The London School of Economics
and Political Science

Misallocation of State Capacity?

Torsten Figueiredo Walter

A thesis submitted to the Department of Economics of the London School of Economics for the degree of Doctor of Philosophy. London, September 2018.

Declaration

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that this thesis consists of approximately 17,000 words.

Abstract

This thesis examines the allocation of human capital in the public sector. I build a new global school-level database comprising 1.73 million public primary schools in 86 countries to study the allocation of teachers across schools across countries at different income levels. In line with common wisdom, I find a strong negative correlation between school-level pupil-teacher ratios (PTRs) and the level of income of a country. More strikingly, I document that the within-country variation in PTRs is also higher in lower income countries. This negative correlation between PTR variation and per capita income is also found within countries over time. Cross-country regressions and cross-district regressions within developing countries suggest that teachers may be misallocated across schools in developing countries: aggregate educational attainment and PTR variation are negatively correlated - even after controlling for differences in per capita income and aggregate PTR. I build a theoretical framework to characterize the notion of misallocation and calibrate the model to simulate counterfactual teacher allocations. I find that aggregate gains in grade promotion from teacher reallocation would be substantial in many developing countries. I finish by discussing the causes and implications of my findings. A case study from Zambia points to lack of managerial capacity and weak enforcement of teacher allocation policies as important underlying factors. A comparison of the distribution of health workers across public primary care facilities in Zambia and England suggests that misallocation of public human resources could also be an issue in other public sectors in developing countries.

Acknowledgments

The completion of this thesis has led to the accumulation of enormous debts, to a great many people, that I will probably never be able to repay.

I would like to thank Nava Ashraf, Oriana Bandiera, Gharad Bryan, and Robin Burgess for extensive advice and support. I also thank Shan Aman Rana, Michel Azulai, Clare Balboni, Andres Barrios Fernandez, Kelsey Jack, Niclas Moneke, Sandra Sequeira, Alessandro Sforza, Frederik Thenee, Sven Walter, Guo Xu, Celine Zipfel, and seminar participants at LSE for valuable comments. I thank Dylan Knaggs and Anton Heil for excellent research assistance.

In addition, I am thankful for help with data collection to Shan Aman Rana, Michel Azulai, Diego Battiston, Svetlana Chekmasova, Joshua Chipman, Andreas Ek, Luiza Figueiredo Walter, Amgalan Amgaa Ganbat, Hanwei Huang, Dana Kassem, Ruth Kattumuri, Sevim Kosem, Panos Mavrokonstantis, Niclas Moneke, Chimgee Mongol, Daniel Morales, Abhiroop Mukhopadhyay, Bupe Musonda, Kieu-Trang Nguyen, Vesall Nourani, Tsogsag Nyamdavaa, Thomas O'Keeffe, Ka Phaydanglobriayao, Giulio Schinaia, Alessandro Sforza, Ella Spencer, Yegor Stadny, Junichi Yamasaki, and Celine Zipfel. Special thanks to Benjamin Chibuye, Miljan Sladoje and Twivwe Siwale at IGC Zambia as well as Alberto da Cruz, Novella Maugeri and Jorrit Oppewal at IGC Mozambique for their support of my in-country work. Thanks also to all the helpful employees at the Ministries of Education and Statistical Agencies in the many countries that are part of this thesis, especially to those in Mozambique and Zambia, as well as those at the Zambian Ministry of Health and EQUIP Zambia.

Thanks are also due to the administrative staff at LSE, with special mention to Rhoda Frith, Mark Wilbor, Jane Dickson, Nic Warner and Michael Rose.

Last, but by no means least, my family and friends deserve special thanks for putting up with me with such grace and good humour throughout this long journey.

Financial support from the Economic and Social Research Council and the International Growth Center is gratefully acknowledged.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Human capital is key for the delivery of public services. There would be no schools without teachers, no hospitals without doctors, no police stations without officers. However, a growing body of evidence points to inefficient human capital allocation in the public sector of developing countries. Politically driven recruitment and transfer decisions are common and human resource management capacity is limited (Akhtari et al. 2017, Asim et al. 2017, Beteille 2009, Lemos & Scur 2016, Sharma & Ramachandran 2009). This raises the question to what extent human capital is misallocated in developing countries. Echoing the literature on capital misallocation across firms (Hsieh & Klenow 2009)[1], this thesis asks whether human capital is misallocated across public institutions.

Focusing on public primary education, I conduct a comparative analysis across 86 countries and ask whether teachers are misallocated across public primary schools within countries. I concentrate on the public primary school sector for four reasons. First, education is paramount for development. Second, public primary schools are arguably one of the most common public institutions worldwide. After all, primary education is nearly universally free and compulsory[2], and by and large

[1] See Restuccia & Rogerson (2017) for a recent review of the literature.

[2] Figures A.1 and A.2 show that primary education is free in 185 countries and compulsory in 195 countries.

publicly provided[3]. Third, teachers are the main input to primary education, as reflected by the large share of government education expenditure going to teacher compensation throughout the world. Across countries, on average 63.5% of total government expenditure for public primary education is spent on teacher compensation[4]. Finally, teachers typically represent a large share of civil servants in a country, with a majority employed in primary education.

To study the allocation of teachers across public primary schools, I build a new global school-level data set comprising 1.73 million public primary schools in 86 countries across all continents and income levels. This data is largely collected from governmental Education Management Information Systems comprising the universe of public primary schools in 70 countries. It is complemented by data from state-level school censuses in six countries and nationally representative surveys from ten countries. In line with common wisdom, I document a strong negative correlation between school-level pupil-teacher ratios (PTRs) and the level of income of a country[5]. More strikingly, I find that the within-country variation in PTRs is also higher in lower income countries. For example, the difference between the 90th and the 10th percentile of the cross-school PTR distribution in Mozambique is 69.7. In the UK, the equivalent difference is as small as 10. This negative correlation between PTR variation and per capita income is also found within countries over time. Cross-country regressions and cross-district regressions within developing countries suggest that teachers may be misallocated across schools in developing countries: aggregate educational attainment and PTR variation are negatively correlated - even after controlling for differences in per capita income and aggregate PTR. The conjecture

[3]Figure A.3 shows the share of primary school pupils enrolled in public institutions in 128 countries. 88.5% of all primary school pupils in these countries attend public institutions.

[4]Figure A.4 shows teacher compensation as a share of total government expenditure on public primary education by country.

[5]This relationship can also be found in publicly available data from the UNESCO Institute for Statistics and World Bank International Comparison Program Database. See figure A.5.

of misallocation is further supported by a third new stylized fact. School remoteness can only explain a small share of PTR variation in developing countries. Thus, the observed PTR differences are not simply the consequence of differences in teacher labor supply or demand for education between rural and urban areas.

To understand to what extent teachers are misallocated across schools in developing countries and how large gains in aggregate education outcomes from teacher reallocation could be, I build a theoretical framework and calibrate it to simulate counterfactual teacher allocations. I consider three counterfactual scenarios. First, teachers are allocated across schools according to a rule that sets a maximum school-level PTR. In each country, this maximum is chosen such that it is the smallest maximum that can be achieved given the distribution of pupils across schools and the total stock of teachers. Second, teachers are optimally allocated across schools. Third, teachers are optimally allocated within subnational administrative units, but cannot be reallocated across units. I simulate these counterfactuals for 20 countries and find that aggregate gains in grade promotion from teacher reallocation in developed countries would be small. Gains in many developing countries in South Asia and Sub-Saharan Africa, on the contrary, would be substantial. For example, estimated gains from the implementation of the smallest achievable maximum PTR rule range from approximately 1 percentage point in Zambia to almost 4 percentage points in India. I compute that these promotion gains imply an additional year of schooling for 1-6% of primary school-aged children in these countries by the time they turn 22. Achieving equivalent gains through the recruitment of additional teachers holding fixed relative PTRs between schools would require teacher workforce increases between 6% and 40%, associated with annual wage costs that range between 1% and 6% of total government education expenditure. Gains from the other two counterfactuals, optimal allocation of teachers and optimal allocation within subnational administrative units, are estimated to be even larger. Thus, teachers appear to be seriously misallocated across public primary schools in developing countries in South Asia and Sub-Saharan Africa.

I finish by discussing the causes and implications of my findings. Given the large estimated gains from teacher reallocation, one has to ask how large the costs of reallocation would be. In order to understand these costs, the causes of teacher misallocation need to be identified in the first place. Therefore, I conduct a case study in Zambia. In collaboration with the Zambian Ministry of General Education, I examine budgeting, deployment, and transfers of teachers. The resulting evidence points to lack of managerial capacity and weak enforcement of teacher allocation policies as important underlying factors of PTR variation. Evidence from other countries (e.g. Asim et al. 2017, Ramachandran et al. 2018) paints a similar picture.

These causes are unlikely to be confined to the public education sector. Other public sectors in developing countries, such as health, law enforcement, and administration, are likely to be similarly affected and therefore human capital misallocation could be an important issue in these as well. In fact, the documented stylized facts on the distribution of public primary school teachers appear to hold in the public health sector as well. Comparing the distribution of health workers across public primary care facilities in Zambia and England, I find a significantly larger dispersion in population per health worker in Zambia than in England. Additionally, only a very small share of the differences in staffing levels across Zambian facilities can be explained by facility remoteness.

This thesis relates to four distinct but interrelated streams of literature. First, it joins a small nascent literature on the allocation of teachers across schools within countries. Work by international organizations such as the World Bank, UNESCO, and UNICEF has repeatedly drawn attention to imbalances in school staffing across districts and schools within specific African countries over many years (e.g. IIEP Pole de Dakar 2016, Mingat et al. 2003, Mulkeen 2010, UNESCO 2006). In recent years, awareness of these imbalances has increased and a few studies have examined implications and causes in selected countries (Agarwal et al. 2016 and Pelkonen & Fagernas 2017 in India, Asim et al. 2017 in Malawi). This thesis focuses on public primary schools and adds a detailed systematic global cross-country comparison to this literature.

Second, this thesis contributes to a long literature on the supply of education in developing countries. This branch of literature has examined to what extent the lack of specific resources - from flip charts (Glewwe et al. 2004) to teacher knowledge (Bold et al. 2017), school monitoring (Duflo et al. 2012, Muralidharan et al. 2017), management practices (Lemos & Scur 2016), and beyond - hampers educational performance. Rather than focusing on the effect of the lack of a given resource, this thesis points to inefficient resource allocation as a hindrance to better aggregate educational outcomes.

Third, it relates to the literature on the personnel economics of the state. While the focus of this literature has been on the selection of state-employees (Ashraf et al. 2018, Dal Bo et al. 2013) and performance incentives (De Ree et al. 2018, Muralidharan & Sundararaman 2011), the long tail of poorly staffed schools in many developing countries revealed in this thesis serves as a reminder that performance-enhancing selection and incentive schemes can only have a limited effect at such schools - unless they also address the prevailing low staffing levels themselves.

Finally, it contributes to the literature on misallocation of production factors across producers in two ways. First, it extends it into the public sector and second, it looks specifically at human captial. This is unlike most of the existing literature which has focused on the misallocation of capital across firms (e.g. Busso et al. 2013, Hsieh & Klenow 2009).

This thesis proceeds as follows. Chapter 2 outlines a simple theoretical framework to characterize the notion of teacher misallocation across schools. In chapter 3, I first describe the data collection process and the resulting data set. Then I document the above mentioned three new stylized facts suggesting that teachers are misallocated. In chapter 4, I calibrate the model sketched in chapter 2 and simulate counterfactual teacher allocations. Chapter 5 discusses the causes of teacher misallocation presenting a case study from Zambia. In chapter 6, I analyse the distribution of health workers across primary care facilities in England and Zambia. Finally, chapter 7 concludes.

Chapter 2

Theoretical framework for the analysis of teacher allocation

In this chapter, I briefly outline a theoretical framework to characterize the notion of teacher misallocation underlying the empirical analysis in chapter 3.

The setup is as follows. A social planner allocates teachers across public primary schools subject to a budget constraint. Teachers are assumed to be homogeneous. This is for two reasons. First, there is typically no subject specialization among primary school teachers. Teachers usually teach all subjects to a given class. Second, differences in teacher quality are not considered due to a lack of available data[1]. The objective of the social planner is universal primary education[2]. As the 2018 World Development Report states, ”most children today enroll in primary school” (World Bank 2018), even in low-income countries[3]. Therefore, I take initial enrollment of pupils as given and focus on school

[1]Teacher quality is typically estimated through teacher value-added models. These require rich panel data on teachers and the performance of their pupils which is not available in most developing countries. Teacher qualifications are more easily available, but it is unclear in how far qualifications actually predict teacher performance.

[2]This is a common goal among policy makers which has been widely promoted by international organizations. It is part of the fourth UN sustainable development goal to ”ensure that all girls and boys complete free, equitable and quality primary and secondary education”.

[3]See figure B.1 for empirical evidence.

completion conditional on initial enrollment[4]. The planner's optimization problem can then be written as follows:

$$\begin{aligned} \max_{T_s} \quad & \sum_s \frac{P_s}{\sum_j P_j} H_s(t_s, .) \\ \text{s.t.} \quad & \sum_s w_s T_s \leq B \end{aligned}$$

where P_s indicates the number of pupils in school s, T_s the number of teachers, and $t_s = T_s/P_s$ is defined as the number of teachers per pupil. H_s is the grade promotion rate, i.e. the share of pupils that advances to the next grade. The grade promotion rate at each school is weighted by the enrollment share of the school $P_s/\sum_j P_j$. Hence, the social planner maximizes the total number of pupils in public primary schools advancing to the next grade. The cost of having a teacher at school s is w_s and the budget constraint says that the total cost for teachers cannot exceed the budget B.

I assume that the grade promotion rate H is a function of school quality $\widetilde{Q}$ and location-specific demand factors L:

$$H = h\left(\widetilde{Q}, L\right)$$

$\widetilde{Q}$ captures school-specific inputs such as the number of teachers per pupil t, learning materials, school infrastructure and school management. L captures location-specific determinants of demand for primary education, such as location-specific household preferences, returns to education, and opportunity costs. h describes the household education investment decision given $\widetilde{Q}$ and L, and will henceforth be referred to as the policy function[5].

[4]Figure B.2 shows that there is substantial variation in primary school completion across countries.

[5]See discussion in Glewwe & Muralidharan (2016).

Given this setup, a necessary condition for an interior solution to the social planner's problem is

$$\frac{\frac{\partial h}{\partial t}(t_i, Q_i, L_i)}{w_i} = \frac{\frac{\partial h}{\partial t}(t_j, Q_j, L_j)}{w_j} \tag{2.1}$$

where $\widetilde{Q} = \{t, Q\}$. This condition illustrates that in this framework cross-school variation in three factors can rationalize PTR variation across public primary schools:

1. School quality (t_i and Q_i)

2. Location-specific demand factors (L_i)

3. Teacher costs (w_i)

Consequently, larger variation in these factors in developing countries than in developed countries could rationalize larger observed PTR variation in the former. In the next chapter, I examine this empirically in order to determine to what extent teachers are indeed misallocated across public primary schools in developing countries.

Chapter 3

Cross-country comparison of the distribution of teachers across public primary schools

3.1 Data

In order to examine the distribution of teachers across public primary schools, I collected school-level data on pupil-teacher ratios for the universe of public primary schools in 70 countries. This data is supplemented by state-level school censuses from six countries and nationally representative school- and household survey data from an additional ten countries. Below I describe the data collection process and the resulting data set.

3.1.1 Data collection

Data collection was carried out in three steps as detailed below.

First, I visited the website of the Ministry of Education of every country in the world to look for school census data. If a Ministry of Education did not have a website or I could not find school census data on their website, I visited the website of the Central Statistical Agency. In countries with a decentralized administration of the education system (e.g. Canada) I also visited websites of subnational education authorities.

This way I found publicly available school census data online for 47 countries.

In a second step, I sent a data request letter to the Ministry of Education and/or Central Statistical Agency of all remaining countries as long as a point of contact (email address or personal contact) could be found. In some countries with decentralized education systems, data requests were sent to state- or province-level authorities. Overall, I sent out more than 250 data requests in five different languages and collected data from 39 countries this way.

Third, for all countries where neither of the two previous approaches had been successful, I checked the availability of nationally representative school survey data with information on school-level PTRs. This way, data for another ten countries was added.

Overall, I obtained data from 86 countries in 14 different languages and many different formats. Finally, I synchronized language and format of the data across countries. Table C.1 gives a detailed overview of the all the data sources and the following subsection provides a description of the resulting data set.

3.1.2 Core data

The final data set contains school-level PTR data from 86 different countries across all continents and income levels. As previously mentioned, countries can be subdivided into three categories. First, for 70 countries, school census data for the national universe of public primary schools was obtained[1]. Second, for six countries, school census data was only obtained from a subset of states or provinces (covering the universe of public primary schools within those). Third, nationally representative

[1]It is difficult to assess whether the school census data indeed covers all public primary schools in each of the sample countries and states. However, data on school census return rates from public schools across 49 African countries from the UNESCO Institute for Statistics and World Bank International Comparison Program Database suggests that even in low-income countries data is fairly complete. Return rates are on average 97.3% and only in a handful of countries they are below 90%. See figure C.1.

survey data was collected for ten countries. Figure 3.1 provides an overview of the geographical coverage of the data set.

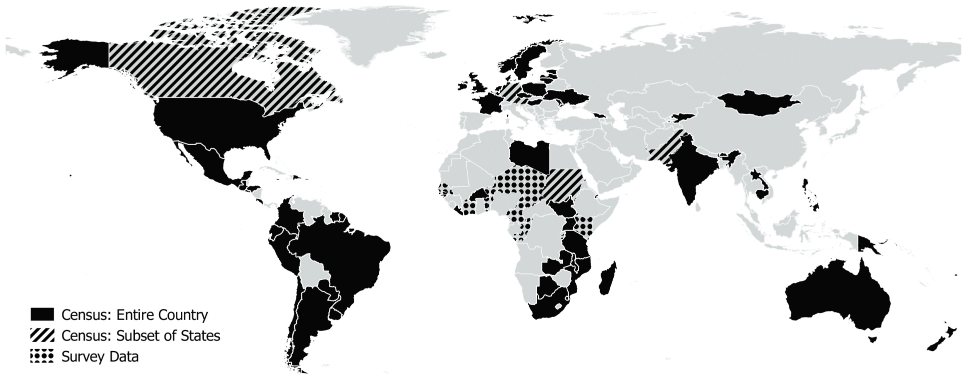


Figure 3.1: Data coverage

The final data set contains information on more than 1.73 million public primary schools attended by a total of approximately 273 million primary school pupils. Given a total world population between the age of 5 and 14 of roughly 1.24 billion in 2015, this means it covers about 22% of all primary school pupils worldwide. The total number of teachers working at these schools adds up to almost 12 million.

For each country, the year of the data corresponds to the latest available year at the time of data collection. The majority of the data is from the time period between 2013 and 2017. Only 7 out of 86 country-level data sets are from before 2013, with the earliest data from Botswana in 2009[2]. Table 3.1 specifies the year of the data for each country and provides basic summary statistics for each country. Further details are documented in table C.1.

PTRs are generally computed as the ratio of pupil headcount over teacher headcount[3]. While it would be desirable to use full-time equivalents instead of headcounts, such data is rarely available. Hence, for the sake of comparability headcounts are used whenever possible.

[2]Figure C.2 shows the number of data sets by year.

[3]Note that school-level teacher headcounts imply that teachers are counted repeatedly if they work in several schools simultaneously.

However, in ten countries teacher headcounts were not available. In seven of these, teacher full-time equivalents could be obtained instead[4]. In the remaining three countries, the total number of school staff (teachers plus management/administration personnel) was used as the denominator[5].

While the data is generally restricted to public primary schools, in five countries the data also contains private primary schools as these cannot be differentiated from public schools in the source data. However, in all of these countries the number of private primary schools is negligibly small[6].

The age at which children start primary education varies little across countries, and is always between 5 and 7[7]. Primary education is most commonly provided through primary schools, but in some countries other school types also provide primary education. For example, in Mongolia primary education is mainly provided at comprehensive schools that run from grade 1 to 12. In order to maintain comparability across schools within each country, I restrict the data to the types of school that are the primary providers of primary education. Table 3.1 lists the included types of schools for each country[8]. The number of grades taught in these schools varies substantially across countries and is also indicated in the table. It reaches from 4 to 12 grades. Primary schools with 6 grades are the most common type[9].

[4]The respective countries where teacher full-time equivalents are used to compute school-level PTRs are Brazil, Canada, Ireland, Puerto Rico, Sweden, the UK, and the US.

[5]The respective countries where school staff headcounts were used instead of teacher headcounts are Belgium (Flanders), Fiji, and France.

[6]The respective countries where the data also contains private primary schools are Cape Verde (0.97%), Fiji (0.86%), Saint Vincent and the Grenadines (10.57%), Swaziland (1.55%), and Ukraine (0.58%). The percentage of enrolment in primary education in private institutions in 2015 is given in brackets (source: World Bank International Comparison Program Database). The information for Swaziland is from 2014 as 2015 data was not available.

[7]Figure C.3 shows the distribution of primary school entrance age across all sample countries. The underlying data is from 2015 and was extracted from the website of the UNESCO Institute for Statistics on 13/07/2017.

[8]Apart from the indicated grades, schools may also include pre-primary education.

[9]Figure C.4 shows the distribution of the maximum number of grades taught at included school types across sample countries.

The survey data used to supplement the school census data comes from two sources. First, data for nine French-speaking countries in West Africa was obtained from PASEC (Programme d'Analyse des Systemes Educatifs de la Conference des ministres de l'Education des Etats et Gouvernements de la Francophonie). In 2014, PASEC carried out nationally representative surveys of primary schools in these nine countries and collected data on the total number of pupils and teachers in each school through interviews with school principals[10]. I restrict the sample of schools in each country to public primary schools, and use sample weights to construct the distribution of PTRs across public primary schools.

Finally, data for Kenya is from UWEZO 2014. This is a nationally representative household survey to assess learning of children, similar to the well-known ASER survey in India. Households are selected through a three stage sampling process. First, districts are randomly drawn (with equal probability). Then enumeration areas (villages) are picked with selection probability proportional to size, and within each enumeration area households are randomly selected. In addition, a government primary school is surveyed in every sample village. The cross-school PTR distribution is constructed based on the schools contained in the UWEZO sample.

Neither the PASEC nor the UWEZO school surveys were designed to generate a nationally representative picture of the PTR variation across public primary schools in the respective countries. Therefore, the measures of cross-school PTR variation for these ten countries which are derived below should be regarded as tentative. All the subsequent results are robust to dropping these ten countries, and the corresponding results are available in the appendix.

[10]The sample frame included all schools with at least one class in grade 6. The probability of drawing a specific school was proportional to the total number of grade 6 pupils in a school.

3.1.3 Complementary data

Two types of complementary data were collected along with the core data.

First, for a subset of 20 countries, not only the latest available school census, but school census data going as far back as possible was obtained. Table C.2 lists those countries and indicates the years for which data was available.

Second, data on the location of schools was collected. For all but three countries with census data, information on the region in which each school is located was gathered. For 51 countries, subregions were also obtained. A region was defined as the highest administrative division available (e.g. state or province) or the statistical division that is closest to it (e.g. NUTS-2). A subregion was analogously defined as the second highest administrative division available (e.g. district) or the statistical division that is closest to it (e.g. NUTS-3). Table C.3 provides further details including the definition of a region and a subregion used throughout this paper for each country.

In addition, for 51 countries GPS coordinates of schools were gathered. These were either downloaded or requested from the corresponding Ministry of Education. For 34 countries, GPS coordinates were available, for 17 countries school addresses were transformed into GPS coordinates using Google Maps Geocoding API. For few countries, coordinates for all schools could be obtained, but overall the coordinate data is fairly complete. The share of schools for which coordinates are available is on average 94%. Table C.4 provides detailed information on the data source for each country and the completeness of the data[11].

[11]Quality of the data varies across countries. In a few countries, a relatively large share of schools has identical coordinates as at least one other public primary school. Table C.4 provides details. In such cases, there may either be several schools within the same building or coordinates do not reflect the actual location of the school, but rather the centroid of the administrative division within which a school is located. Without additional information, it is not possible to differentiate between these two cases. Therefore, all subsequent results using GPS coordinates are replicated including and excluding schools that share identical coordinates with other schools.

Table 3.1: Data summary

Country	**Data type**	**Year**	**School types (grades)**	**Schools**	**Pupils (K)**	**Teachers (K)**
Antigua and Barbuda	Census	2010/11	PRIM (1-6)	30	5.1	0.4
Argentina	Census	2015	PRIM (1-6, 1-7)	18408	3482.3	273.6
Australia	Census	2016	PRIM (1-6)	4735	1437.3	95.6
Austria	Census*	2016/17	PRIM (1-4)	1771	176.1	20.0
Belgium	Census*	2017	PRIM (1-6)	918	267.4	29.9
Benin	Survey	2014	PRIM (1-6)	133	44.1	0.7
Bhutan	Census	2015	PRIM (1-6)	413	45.4	2.5
Botswana	Census	2009	PRIM (1-7)	707	298.4	11.5
Brazil	Census	2015	PRIM (1-8, 1-9)	79805	22132.3	1040.3
Burkina Faso	Census	2017	PRIM (1-6)	11537	2429.0	58.4
Burundi	Survey	2014	PRIM (1-6)	165	102.8	2.6
Cambodia	Census	2014	PRIM (1-6)	6164	1784.0	40.6
Cameroon	Survey	2014	PRIM (1-6)	177	47.4	0.9
Canada	Census*	2014-16	PRIM (1-5, 1-8)	3471	1326.1	73.7
Cape Verde	Census	2014/15	PRIM (1-6)	407	63.6	2.9
Chad	Survey	2014	PRIM (1-6)	93	42.9	0.7
Chile	Census	2015	PRIM (1-8)	3764	666.2	56.5
Colombia	Census	2015	PRIM (1-5)	39953	3488.8	40.0
Congo, Rep.	Survey	2014	PRIM (1-6)	81	35.0	0.4

Table 3.1: Data summary

Country	Data type	Year	School types (grades)	Schools	Pupils (K)	Teachers (K)
Costa Rica	Census	2016	PRIM (1-6)	3674	538.7	28.6
Cote d'Ivoire	Survey	2014	PRIM (1-6)	144	40.9	0.8
Czech Republic	Census	2017	PRIM (1-9)	3900	886.8	70.4
Denmark	Census	2015/16	COMP (1-10, 1-11)	1204	524.6	55.0
Djibouti	Census	2014/15	PRIM (1-5)	132	56.3	1.6
Dominican Republic	Census	2016/17	PRIM (1-6)	3936	777.4	40.7
Ecuador	Census	2015/16	PRIM (1-7)	9245	790.1	33.1
El Salvador	Census	2013	PRIM (1-9)	4224	912.1	30.1
Estonia	Census	2016	PRIM (1-3, 1-6)	415	79.7	10.5
Fiji	Census	2017	PRIM (1-6)	691	142.3	5.7
France	Census	2015/16	PRIM (1-5)	29550	3930.9	189.4
Georgia	Census	2016	COMP (1-12)	1587	486.7	53.1
Germany	Census*	2014/15	PRIM (1-4)	2017	389.8	28.5
Guatemala	Census	2015	PRIM (1-3, 1-6, 4-6)	19448	2453.2	108.4
Guinea-Bissau	Census	2014	PRIM (1-4, 1-6)	694	170.7	5.7
Honduras	Census	2012	PRIM (1-6, 1-9)	11440	1192.8	79.4
Hungary	Census	2015/16	PRIM (1-8)	2952	629.4	65.0
India	Census	2015	PRIM (1-5, 6-8, 1-8)	981351	97279.2	3825.0
Ireland	Census	2015/16	PRIM (1-6)	3124	541.0	31.8

Table 3.1: Data summary

Country	**Data type**	**Year**	**School types (grades)**	**Schools**	**Pupils (K)**	**Teachers (K)**
Jamaica	Census	2015	PRIM (1-6)	377	136.0	4.7
Kenya	Survey	2014	PRIM (1-8)	4135	2373.6	68.6
Kiribati	Census	2011	PRIM (1-6)	94	15.5	0.6
Kyrgyzstan	Census	2015	COMP (1-11)	1674	822.8	60.6
Laos	Census	2016/17	PRIM (1-5)	8606	807.4	38.1
Latvia	Census	2016	PRIM (1-6) & BAS (1-9)	369	51.0	7.4
Liberia	Census	2015	PRIM (1-6) & BAS (1-9)	2486	335.7	12.2
Libya	Census	2012	PRIM (1-6) & BAS (1-9)	3194	1005.4	175.9
Lithuania	Census	2016	PRIM (1-4) & BAS (1-10)	555	57.3	4.5
Madagascar	Census	2016	PRIM (1-5)	24447	3857.8	89.8
Malawi	Census	2016	PRIM (1-8)	5404	4703.5	68.4
Marshall Islands	Census	2013/14	PRIM (1-8)	80	9.7	0.7
Mexico	Census	2015/16	PRIM (1-6)	88991	12969.9	514.0
Moldova	Census	2016	PRIM (1-4) & BAS (1-9)	906	69.3	4.4
Mongolia	Census	2016	BAS (1-9) & COMP (1-12)	614	272.1	8.9
Mozambique	Census	2016	PRIM (1-5, 6-7, 1-7)	12386	5815.3	108.2
Netherlands	Census	2015/16	PRIM (1-8)	2059	437.1	36.2
New Zealand	Census	2015	PRIM (1-5, 1-7, 6-7)	1691	405.1	26.1
Niger	Survey	2014	PRIM (1-6)	166	61.8	1.5

Table 3.1: Data summary

Country	Data type	Year	School types (grades)	Schools	Pupils (K)	Teachers (K)
Norway	Census	2016/17	PRIM (1-7)	2114	429.4	40.6
Pakistan	Census*	2013-16	PRIM (1-5)	80593	7044.5	206.6
Palau	Census	2016	PRIM (1-6)	18	1.8	0.2
Papua New Guinea	Census	2016	PRIM (1-6)	4264	732.9	19.8
Paraguay	Census	2013/14	PRIM (1-6, 1-9)	5176	676.2	62.3
Peru	Census	2016	PRIM (1-6)	29141	2563.5	140.6
Philippines	Census	2013/14	PRIM (1-6)	37948	14952.8	377.5
Poland	Census	2017	PRIM (1-6)	9577	2675.3	331.1
Puerto Rico	Census	2014/15	PRIM (1-5, 1-6, 1-7, 1-8)	771	180.8	14.0
St Kitts and Nevis	Census	2013/14	PRIM (1-6)	24	4.3	0.3
St Lucia	Census	2014/15	PRIM (1-9)	74	15.8	1.0
St Vincent and the Grenadines	Census	2014/15	PRIM (1-6)	68	13.4	0.9
Samoa	Census	2015	PRIM (1-8)	143	33.7	1.1
Senegal	Survey	2014	PRIM (1-6)	134	68.6	1.4
Seychelles	Census	2012	PRIM (1-6)	24	10.4	0.9
South Africa	Census	2015	PRIM (1-7)	13781	6497.8	190.8
South Sudan	Census	2015	PRIM (1-8)	2409	884.7	20.2
Sudan	Census*	2012	PRIM (1-8)	1309	498.0	17.1
Suriname	Census	2016	PRIM (1-7)	333	69.6	5.9

Table 3.1: Data summary

Country	Data type	Year	School types (grades)	Schools	Pupils (K)	Teachers (K)
Swaziland	Census	2013	PRIM (1-7)	591	235.2	8.2
Sweden	Census	2015/16	PRIM (1-9)	3982	838.4	68.8
Tanzania	Census	2016	PRIM (1-7)	14598	7489.3	172.5
Togo	Survey	2014	PRIM (1-6)	141	35.9	0.7
UK	Census	2015/16	PRIM (1-4)	20118	5289.6	255.1
US	Census	2014/15	PRIM (1-5, 1-6, 1-7, 1-8)	51732	24046.2	1471.0
Uganda	Census	2016	PRIM (1-7)	11357	6702.4	122.4
Ukraine	Census	2013/14	COMP (1-10, 1-11)	16370	3771.4	443.0
Uruguay	Census	2015	PRIM (1-6)	1953	247.8	11.4
Zambia	Census	2015	PRIM (1-7) & BAS (1-9)	5790	2864.1	62.0

The table indicates three different types of data: census, census*, and survey. Census means that data for the universe public primary schools was collected. Census* indicates that data for the universe of public primary schools was collected from a subset of the highest administrative divisions in the country. See table C.1 for details. Survey indicates that the collected data is from a nationally representative school or household survey. Column (4) lists the included school types for each country and the typical grade range at these schools. PRIM stands for primary, BAS for basic, and COMP for comprehensive. Apart from the indicated grades, schools may also include pre-primary education. The last three columns provide information on the total number of schools, pupils and teachers contained in the data. Totals are computed after dropping schools for which PTR information was not available. See table C.1 for details on the share of public primary schools without PTR information.

3.2 Stylized facts

In this section, I compare the distribution of teachers across public primary schools between and within countries based on the assembled data. I document three new stylized facts that are jointly suggestive of teacher misallocation across public primary schools in developing countries. These facts are:

1. PTR variation is negatively correlated with per capita income, both across countries and within countries across time.
2. PTR variation is negatively correlated with aggregate education outcomes even after controlling for differences in income and aggregate PTRs.
3. In low- and lower-middle-income countries, PTRs are larger in remote schools, but remoteness can only explain a small share of overall PTR variation.

The first two facts on their own are suggestive of misallocation. The third fact rules out that PTR variation in developing countries can be rationalized by spatial variation in teacher labor supply or demand for education between rural and urban areas.

3.2.1 Fact 1

PTR variation is negatively correlated with per capita income, both across countries and within countries across time.

Figure 3.2 shows the PTR distribution across public primary schools in four different countries - a low-income country (Mozambique), a lower-middle-income country (India), an upper-middle income country (Peru) and a high-income country (UK) - in black[12]. Grey dashed lines represent

[12]The data for each country are trimmed at bottom and the top. The 1st and the 99th percentile of the PTR distribution are excluded. All the following results are robust to including these.

the enrollment-weighted PTR distribution across public primary schools. A comparison of these four distributions yields three observations. First, average PTR in the low-income country Mozambique is substantially higher than in the other countries. While the mean PTR is 59.8 in Mozambique, it is only 24.9 in India, 14.2 in Peru and 20.4 in the UK. Second, cross-school PTR variation is large in the lower income countries, but small in the higher income countries. The cross-school PTR standard deviation amounts to 27.4 and 18.2, respectively, in Mozambique and India, but it measures only 6.8 and 3.8, respectively, in Peru and the UK[13]. Third, the poorer a country, the longer is the right tail of schools with high PTRs. These three cross-country observations hold independently of whether schools are weighted by their enrollment or not. They also hold across countries more generally and I document this below.

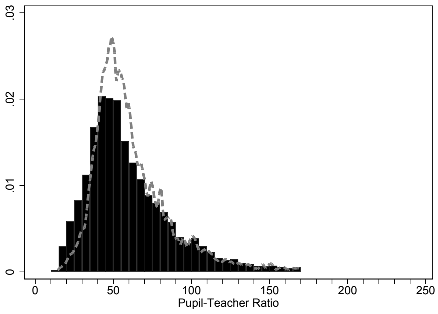
As outlined in the introduction, it is known that aggregate PTRs in primary education are higher in low income-countries[14]. Since the large majority of primary education is publicly provided, it is not surprising that figure 3.3 confirms this relationship for the public primary education sector[15].

Figure 3.4 illustrates the negative relationship between cross-school PTR variation in the public primary education sector and per capita income across countries. It plots the cross-school PTR standard deviation within each country against per capita income. Note that the negative association between PTR dispersion and per capita income is not simply a consequence of high aggregate PTRs in low-income countries. While it can be shown through simulations that the indivisibility of teachers causes PTR variation to increase in aggregate PTR even if the objective is

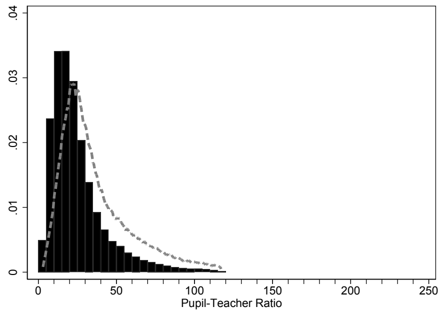
[13]All figures are based on the unweighted distributions illustrated in black.

[14]See figure A.5.

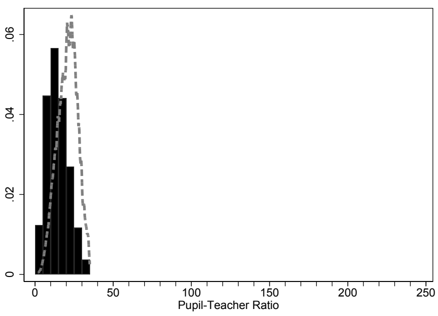
[15]For a given country, the national aggregate PTR is computed as the total number of pupils over the total number of teachers in all public primary schools contained in the data. The total number of teachers is computed as the sum of teacher headcounts over all schools. To the extent that teachers work in several schools simultaneously and are counted repeatedly, computed national aggregate PTRs will underestimate actual national aggregate PTRs. Per capita income data comes from the World Bank International Comparison Program Database.



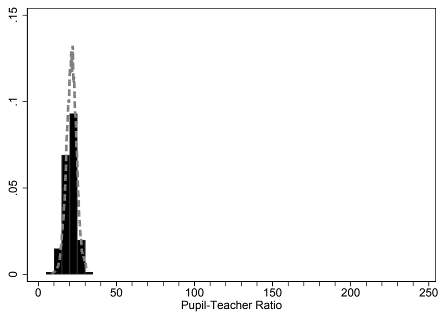
(a) Mozambique



(b) India



(c) Peru



(d) UK

Figure 3.2: Distribution of PTRs across public primary schools in selected countries

Histograms show the distribution of PTRs across the universe of public primary schools in the selected countries. The grey dashed line in each subfigure indicates the cross-school PTR distribution when schools are weighted by their enrollment.

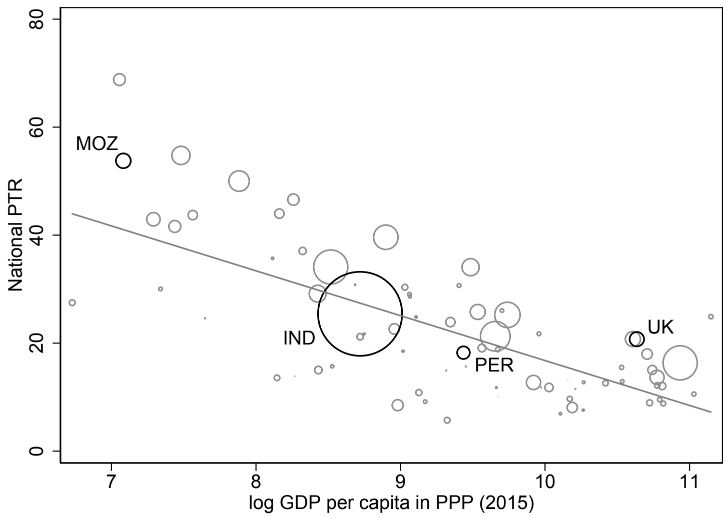


Figure 3.3: National PTR in public primary education and income across countries

The national PTR in public primary education is defined as the total number of pupils in public primary schools over the total number of teachers in these schools. Since these totals are not available for the ten countries with survey data, these countries are not included in the figure. Marker size indicates the size of the primary school-aged population (ages 5 to 14) in a country. GDP per capita and population data are from the World Bank International Comparison Program database.

to equalize PTRs across schools, this effect is quantitatively small relative to the PTR variation observed in low- and middle-income countries[16]. Moreover, the relationship between PTR variation and per capita income remains significantly negative when the coefficient of variation is used as an alternative measure of PTR variation[17].

Finally, figure 3.5 confirms the earlier observation that there is a long right tail of schools with high PTRs in lower income countries which does not exist in higher income countries. The length of the tail of the distribution is measured by the difference between the 90th and the 50th percentile of the cross-school PTR distribution. The long right tail in many developing countries implies that a lot of children attend schools with few teachers - even in countries where national PTRs in the public

[16]See appendix section **??** for details.

[17]See figure C.7.

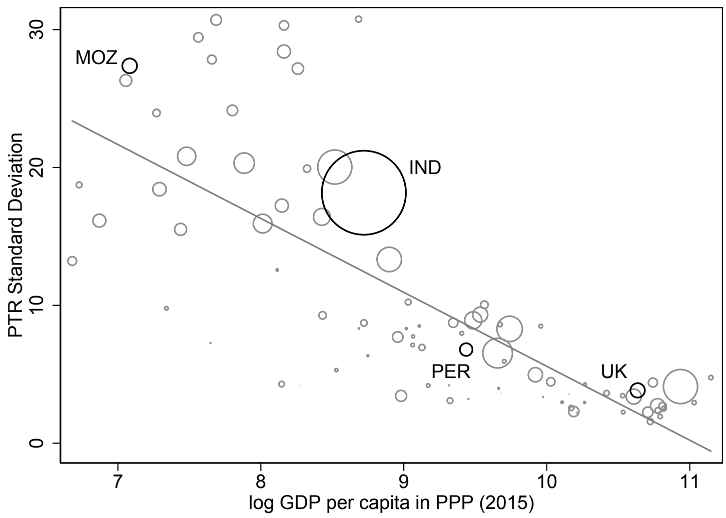


Figure 3.4: PTR variation in public primary education and income across countries

The PTR standard deviation is defined as the standard deviation in PTRs across all public primary schools within a country. The grey line is a linear regression line. Marker size indicates the size of the primary school-aged population (ages 5 to 14) in a country. GDP per capita and population data are from the World Bank International Comparison Program database.

primary education sector are not extremely high. In India, for example, 34.2% of public primary education pupils attend schools with a PTR above 40 despite a national PTR of 24.9. Across all sample countries with census data, 5.6% of children (about 15 million) are enrolled in schools with PTRs above 80 while this share could be reduced to zero if teachers were more evenly distributed across public primary schools within countries.

The documented relationship between per capita income and cross-school PTR variation in the public primary education sector does not only hold across countries but can also be observed within countries across time. This suggests that the cross-country relationship is not merely driven by time-invariant differences between countries that are correlated with income (e.g. geography). Figure 3.6 shows the combined time series of the cross-school PTR standard deviation and per capita income for the 20 countries where such data was collected. It stands out that at low levels

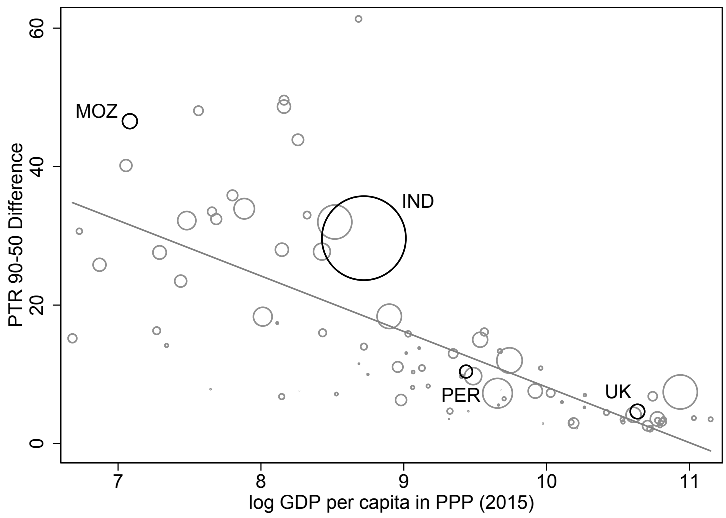


Figure 3.5: Right tail of the cross-school PTR distribution and income across countries
The PTR 90-50 difference is defined as the difference between the 90th percentile and the median of the PTR distribution across all public primary schools within a country. Marker size indicates the size of the primary school-aged population (ages 5 to 14) in a country. GDP per capita and population data are from the World Bank International Comparison Program database.

of GDP per capita there is a strong negative relationship between per capita income and PTR variation whereas there is hardly any relationship at high income levels.

3.2.2 Fact 2

PTR variation is negatively correlated with aggregate education outcomes even after controlling for differences in income and aggregate PTRs.

Table 3.2 shows that across countries, aggregate primary school attainment is negatively correlated with PTR variation, even after controlling for differences in per capita income, population and aggregate PTR. An increase in the within-country PTR standard deviation from 25th to the 75th percentile in the cross-country distribution is associated with an 8.43 (7.03) percentage point decrease in the primary school completion

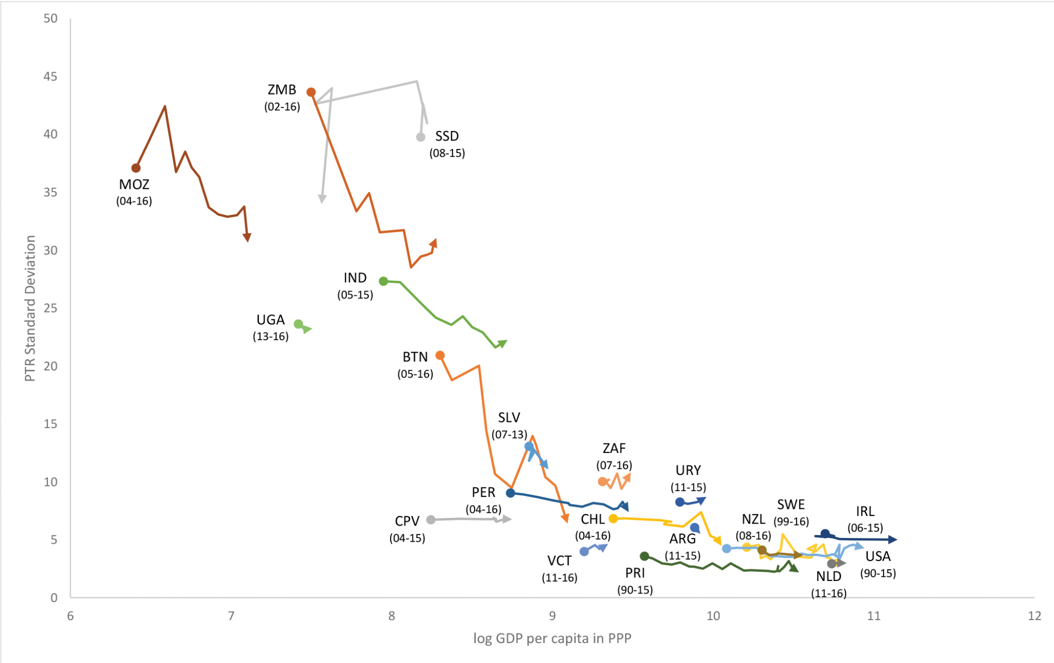


Figure 3.6: PTR variation in public primary education and income over time across countries

The figure illustrates the joint path of the cross-school PTR standard deviation in public primary education and per capita income for 20 countries. Country codes indicate which time series represents which country. The time span covered by the underlying time series data is given in brackets. The starting point of each time series is marked by a dot and the end point by an arrow. The PTR standard deviation is defined as the standard deviation in PTRs across all public primary schools within a country in a given year. GDP per capita data is from the World Bank International Comparison Program database.

(survival) rate[18]. Moreover, columns (3) and (6) indicate that the negative partial correlation between the PTR standard deviation and primary schooling outcomes is driven by lower income countries whereas there is no significant correlation for higher income countries.

While these correlations are intriguing, they could entirely be driven by omitted country characteristics, such as institutional quality. To alleviate this concern to a certain extent, I replicate the above cross-country analysis comparing subregions (districts) within specific low- and middle-income countries. This is feasible because a substantial share of PTR variation in developing countries is within subregions. Figure 3.7 shows the spatial variation of PTRs across public primary schools in Zambia. The map indicates areas around schools with high PTRs in increasingly dark shades of red and areas around school with low PTRs with increasingly dark shades of green. It stands out that the heat map is relatively spotty, i.e. there is a lot of variation even within districts. A similar pattern can be observed in other developing countries[19]. A PTR variance decomposition shows that both between- and within-region and -subregion variation are larger in lower income countries. But while within- and between-variation are of similar magnitude in high-income countries, in lower income countries the within-variation is substantially larger than the between-variation[20].

Cross-district regressions in Mozambique, India, and Peru reveal a similar pattern as the above cross-country reression (see table 3.3). In all three countries, there is a significant negative correlation between public primary schooling outcomes and cross-school PTR variation within

[18]The primary school survival rate is defined as the survival rate until the last grade of primary education. Outcome variables are from the UNESCO Institute for Statistics and World Bank International Comparison Program Database. See *http://data.uis.unesco.org/* for details. For each country the most recent available data as of 06/05/2017 is used. Note that data is not available for all sample countries. In addition, countries with survey data are excluded from the regression because the national aggregate PTR is not available for these countries.

[19]PTR heat maps from other countries where school coordinates were obtained are available upon request.

[20]See figures C.10 and C.11.

Table 3.2: PTR variation and primary schooling outcomes across countries

	Primary school completion			Primary school survival		
	(1)	(2)	(3)	(4)	(5)	(6)
National PTR	-0.00476**	-0.000242	-0.000681	-0.00556***	-0.00196	-0.00213
	(0.00199)	(0.00265)	(0.00270)	(0.00146)	(0.00201)	(0.00197)
PTR SD		-0.0128**			-0.0107**	
		(0.00518)			(0.00428)	
(PTR SD)x(GDP pc high)			-0.00662			-0.00243
			(0.00869)			(0.00625)
(PTR SD)x(GDP pc low)			-0.0122**			-0.0108**
			(0.00522)			(0.00419)
log GDP pc	0.0493**	0.0257	0.0162	0.0723***	0.0529***	0.0382*
	(0.0239)	(0.0249)	(0.0271)	(0.0187)	(0.0195)	(0.0208)
log Population	-0.00117	0.00525	0.00461	-0.00989	-0.00322	-0.00370
	(0.00792)	(0.00806)	(0.00810)	(0.00643)	(0.00670)	(0.00657)
R2	0.418	0.470	0.477	0.730	0.759	0.773
N	67	67	67	57	57	57
Mean Dep. Var.	0.913	0.913	0.913	0.837	0.837	0.837
PTR SD IQR	6.598	6.598	6.598	6.598	6.598	6.598

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The primary school survival rate is defined as the survival rate until the last grade of primary education. The latest available data as of 06/05/2017 is used for each country. "Log GDP pc" is the logarithm of the GDP per capita in PPP terms. "GDP pc low (high)" is a dummy variable that takes value one when the logarithm of the GDP per capita in PPP terms is below (above) the median. Outcome variables and income and population data (both 2015) are from the UNESCO Institute for Statistics and World Bank International Comparison Program Database. The national PTR is defined as the ratio of the total number of public primary school pupils over the total number of public primary school teachers in a country. The PTR SD is defined as the PTR standard deviation across all public primary schools in a country. PTR SD IQR indicates the interquartile range in the PTR SD across countries. See tables C.5 and C.6 for robustness checks. Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

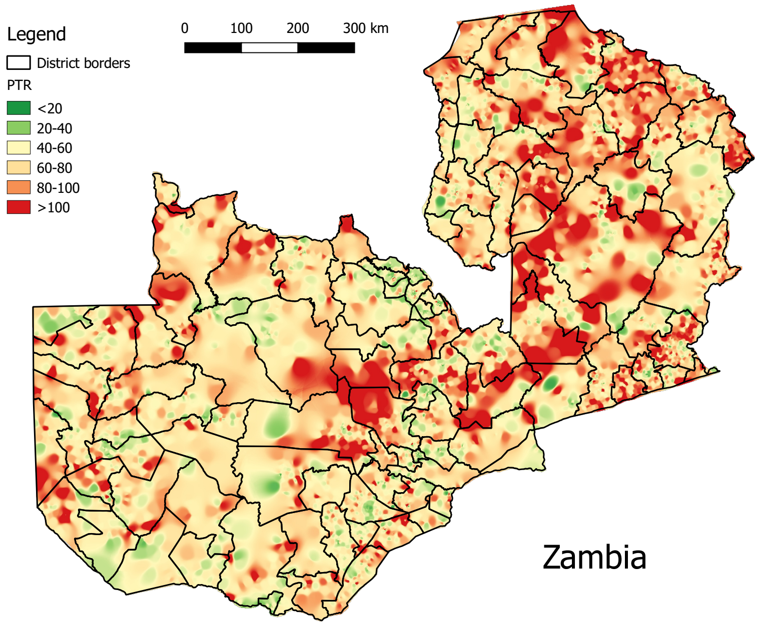


Figure 3.7: PTR heat map Zambia
The heat map shows the spatial variation in PTRs across public primary schools in Zambia. The map indicates areas around schools with high PTRs in increasingly dark shades of red and areas around school with low PTRs with increasingly dark shades of green. The map is based on all public primary schools for which school coordinates were available (see table C.4).

districts[21], even after controlling for differences in income as proxied by nighttime luminosity, population, and aggregate PTR. In Mozambique an increase in the within-district PTR standard deviation from the 25th to the 75th percentile of the cross-district PTR standard deviation distribution is associated with a 3.2 percentage point decrease in the annual grade promotion rate in public primary education. Over 5 years of primary education, this accumulates to a difference of 14 percentage points in completion rates. In India, the pass rate at the national grade 5 exams in the district at the 75th percentile of the PTR standard deviation distribution is 11.5 percentage points lower than the one of the district at the 25th percentile. And even in Peru where PTR dispersion varies

[21] In Peru districts are called provincias.

relatively little across districts and grade completion rates are high, there is a difference in annual grade promotion rates of 0.9 percentage points between the districts at the 75th and the 25th percentile. As in the case of Mozambique, this implies a substantially larger gap in completion rates after 5 years, namely of about 3.8 percentage points.

Table 3.3: PTR variation and public primary schooling outcomes across districts

	Mozambique Grade promotion rate	India Grade 5 exam pass rate	Peru Grade promotion rate
Nighttime Luminosity	0.00656	-0.00563***	0.000228
	(0.0118)	(0.00122)	(0.00101)
log Population	0.0153	0.0277***	0.0101***
	(0.0103)	(0.00725)	(0.00206)
District PTR	-0.000673	0.00565***	-0.00236***
	(0.000867)	(0.00125)	(0.000666)
PTR SD	-0.00238**	-0.0119***	-0.00365**
	(0.000922)	(0.00171)	(0.00154)
R^2	0.106	0.114	0.267
N	142	616	195
Mean Dep. Var.	0.751	0.587	0.956
PTR SD IQR	13.2874	9.6222	2.4872

In Peru, districts are called provincias. Nighttime luminosity is defined as the median nighttime luminosity across grid cells within a district in 2015. VIIRS nighttime lights data is from the Earth Observation Group, NOAA National Geophysical Data Center. Population data for each district is from the latest available census from each country (Mozambique 2017, India 2011, Peru 2007). The district PTR is defined as the ratio of the total number of public primary school pupils over the total number of public primary school teachers in a district. The PTR SD is defined as the PTR standard deviation across all public primary schools in a district. PTR SD IQR indicates the interquartile range in the PTR SD across districts. Table C.1 indicates the sources of the PTR and outcome variables. The promotion completion rates in Mozambique and Peru are defined as the share of public primary school pupils that pass on to the next grade at the end of the school year. The grade 5 exam pass rate in India is defined as the share of public primary school pupils that pass the national grade 5 examination with a score of 60 or above. Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.2.3 Fact 3

In low- and lower-middle-income countries, PTRs are larger in remote schools, but remoteness can only explain a small share of overall PTR variation.

To assess to what extent school remoteness can explain variation in PTRs in developing countries, I construct four distinct measures of school remoteness for all schools for which coordinates were obtained:

1. Population density within a circle of 3km radius around the school based on Global Human Settlement (GHS) data
2. Population density within a circle of 3km radius around the school based on data from the Gridded Population of the World (v4)
3. Nighttime luminosity within a circle of 3km radius around the school based on 2015 data from the Earth Observation Group, NOAA National Geophysical Data Center
4. Travel time to closest city based on the accessibility to cities dataset from the Malaria Atlas Project at Oxford University (Weiss et al. 2018)

Figure 3.8 shows kernel density estimates of the PTR distribution by quartile of school remoteness as measured by population density (GHS) in Mozambique[22]. The figure makes clear how much variation in PTRs there is within each quartile and how little of the variation is explained by remoteness.

To test whether this is also the case in other countries, I run a separate regression of the following form for each country with available school coordinates:

$$PTR_s = \alpha_r + \beta remote_s$$

where α_r stands for a set of region fixed effects and $remote_s$ for the remoteness of school s as measured by one of the four measures listed

[22] Results look very similar for the other three measures of remoteness. See figures C.12, C.13 and C.14

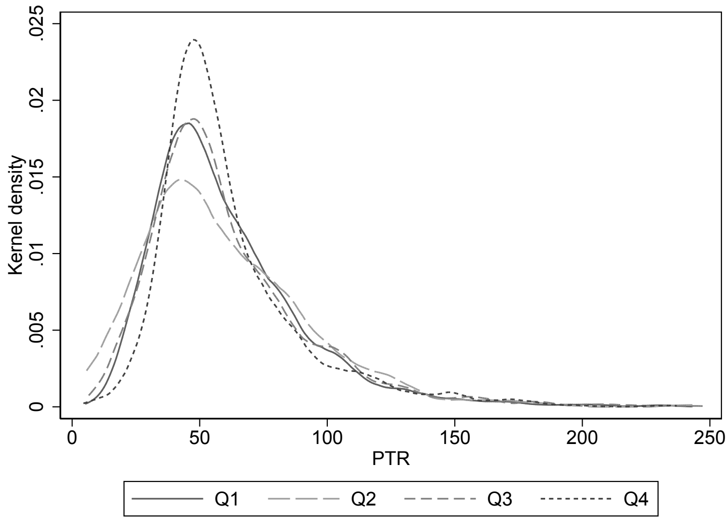
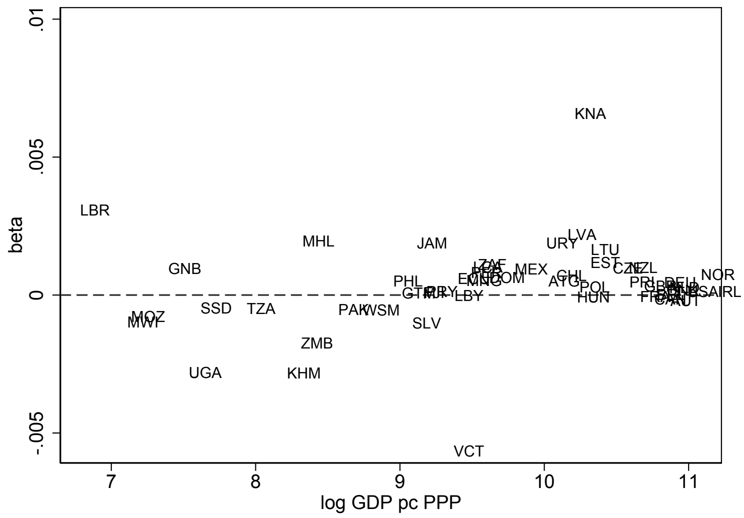


Figure 3.8: PTR distribution by quartile of population density (GHS) in Mozambique
Population density at each public primary school is measured as the density within a circle of 3km radius around the school according to data from the Global Human Settlement project.

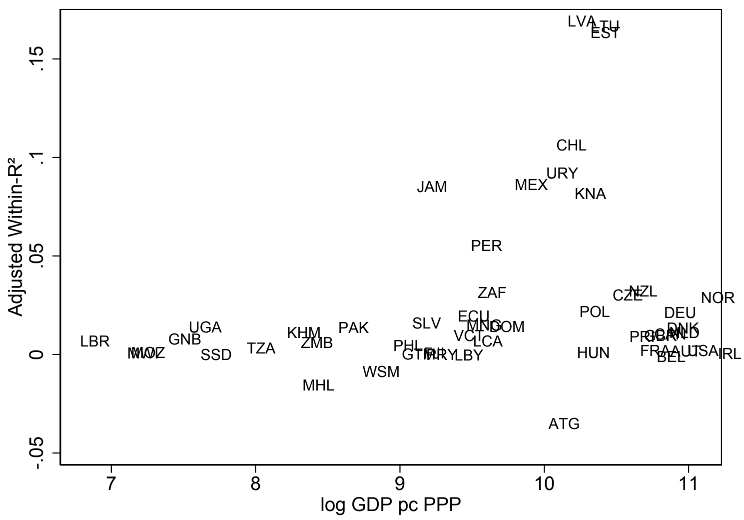
above. Figure 3.9 plots the estimated regression coefficients and the adjusted within-region R^2 for each country against per capita income. I find that remoteness is weakly positively correlated with PTRs in developing countries, but it can only explain a very small share of the overall variation in PTRs as indicated by the low R^2 in these countries[23].

The presented evidence suggests that there could be sizable gains from distributing teachers more equally across schools. Additionally, the fact that a large share of cross-school PTR variation is within subregions means that teacher re-allocation within these could go a long way towards a more balanced distribution. In order to assess how large gains in aggregate education could actually be I proceed by further developing the

[23]The remoteness measure underlying the figure is population density as given by GHS, but results look very similar using the other measures. See figures C.15, C.16 and C.17. A corresponding analysis using rural/urban indicators as provided in the school census data in a subset of 30 countries also leads to similar conclusions. See figure C.18. Since the underlying rural/urban indicators are not comparable across countries, results should be interpreted with care.



(a) Regression coefficient



(b) Within-region R^2

Figure 3.9: PTRs and population density (GHS) across countries by income

Beta is the regression coefficient from a country-specific school-level regression of PTR on population density within a circle of 3km around the school as given by Global Human Settlement data, controlling for region fixed effects. The adjusted within-region R^2 *is from the same regression. Regions are defined as detailed in table C.3. The sample is restricted to 51 countries for which school coordinates were obtained. GDP per capita data is from the World Bank International Comparison Program database.*

initially presented theoretical framework, calibrating the resulting model, and simulating alternative teacher distributions in the following section.

Chapter 4

Gains from Teacher Reallocation? Model Calibration and Simulation

In order to gain an understanding of the potential gains from teacher reallocation and thereby the extent of current misallocation, I simulate three distinct counterfactual scenarios across 20 countries:

1. Smallest achievable maximum PTR rule

2. Optimal allocation

3. Optimal allocation within subnational units

In the first case, I ask how large gains in aggregate grade promotion of public primary school pupils would be if countries distributed teachers according to a rule that sets a maximum school-level PTR. In each country, this maximum is chosen such that it is the smallest maximum that can be achieved given the distribution of pupils across schools and the total stock of teachers. In the second case, I ask what the optimal allocation of teachers would be such that aggregate grade promotion is maximized. The third case restricts teacher reallocation to within subnational units (e.g. states or districts) and examines the optimal allocation under this restriction. I estimate reallocation gains using a calibrated model that originates from the theoretical framework outlined in chapter 2. This model is presented in the next section.

4.1 Model calibration

In chapter 2, the social planner's problem was formalized as follows:

$$\begin{aligned} \max_{T_s} \quad & \sum_s \frac{P_s}{\sum_j P_j} h(t_s, Q_s, L_s) \\ \text{s.t.} \quad & \sum_s w_s T_s \leq B \end{aligned}$$

In order to take this framework to the data, two key elements need to be modelled and calibrated:

1. The effect of school-level PTR on grade promotion $\frac{\partial h}{\partial PTR}(PTR_s, Q_s, L_s)$
2. Relative teacher costs w_i/w_j

Ideally, one would like to understand the precise functional form of $\frac{\partial h}{\partial PTR}(PTR_s, Q_s, L_s)$. As I argue in the next subsection, this is not possible given the limited available empirical evidence. Therefore, I parameterize the policy function h with two parameters. Additionally, I argue that teacher costs are approximately equal between schools.

4.1.1 The effect of PTR on grade promotion

The school-level PTR is likely to affect pupil achievement through at least three channels. The first and most studied one of these is class size. A smaller number of teachers at a given school is likely to go hand in hand with larger class sizes. The second channel is multigrade teaching. Especially in small schools, a lack of teachers is frequently accompanied by multigrade teachhing. The third channel is instruction time. If the number of teachers in a school is not sufficient, multiple shifts can be introduced to avoid overcrowded classrooms. In such cases, instruction time is reduced as the working time of teachers is distributed across shifts. Below I briefly discuss the available empirical evidence on each of these three channels.

The empirical evidence on class-size effects is mixed. Some papers find small negative effects on test scores (e.g. Angrist & Lavy 1999, Krueger 1999), others find no significant effects (e.g. Hoxby 2000, Angrist et al. 2017). It also has to be noticed that most studies are based in developed countries where class sizes rarely exceed 40. Therefore, it is unclear in how far their findings apply to developing country settings. The only existing study from a developing country with large class sizes reports small positive effects from class size reductions (Duflo et al. 2015). However, the authors of this study stress the importance of complementary inputs, especially teacher incentives, for the effectiveness of class size reductions.

Evidence on the other two channels is more clear-cut. Evidence on the effects of multi-grade teaching is limited, but the available research suggests that it is harmful to student performance (Checchi & De Paola 2017, Jacob et al. 2008). The literature on the effects of instruction time on pupil performance finds largely positive effects (e.g. Lavy 2015). However, it also points out that magnitudes depend on the quality of instruction and children's alternative time use (e.g. Rivkin and Schiman 2015)[1].

There is only one paper that directly investigates the effect of school-level PTRs on pupil performance. Muralidharan & Sundararaman (2013) conducted a randomized control trial across public primary schools in rural Andhra Pradesh, India. Treatment schools obtained an extra contract teacher. This induced an average PTR reduction by 10.814 (after two years). The authors show that a one unit reduction in PTR led to an increase in standardized test scores by 0.0144 standard deviations. They do not find any evidence of heterogeneous effects with respect to student and household characteristics[2].

In order to translate these effects on standardized test scores into effects on grade promotion rates, I assume that test scores are normally distributed. In this case, standardized test scores are standard normally distributed. Given that only pupils with sufficiently high test scores are

[1] See the literature section Barrios & Bovini (2017) for a short summary of the literature.

[2] Chin (2005) also finds positive effects from adding teachers to small schools in India on school completion rates, but does not report the induced reductions in PTRs.

promoted to the next grade, I parameterize the policy function as follows:

$$H_s = 1 - \Phi\left(\alpha_s - \beta\frac{P_s}{T_s}\right)$$

where α_s indicates the productivity level of school s and Φ indicates the standard normal cumulative distribution function. Given the current promotion rate c_s and PTR r_s as well as β, I back out α_s:

$$\alpha_s = \Phi^{-1}(1 - c_s) + \beta r_s$$

Muralidharan and Sundararaman (2013) estimate the average effect of treatment-induced PTR reduction after one and two years. After one year the effect amounts to 0.0099, after two years to 0.0144. It is unknown whether the effect increases further in subsequent periods. To understand the long-run effects of teacher reallocation, it would be desirable to set *beta* to its long-run value. But in absence of empirical evidence, I set $\beta = 0.0144$ assuming no further increases in the effect after the second year. Figure 4.1 illustrates the resulting policy functions for a low- and a high-productivity school.

4.1.2 Relative teacher costs

It is well known that teachers prefer working in urban areas (Fagernas & Pelkonen 2012, Sow 2016) and governments of developing countries report problems with teacher recruitment and retention in rural areas. In order to attract teachers to rural schools governments frequently pay hardship allowances (Pugatch & Schroeder 2013). But while school remoteness is positively correlated with PTRs as suggested by this evidence, it can only explain a small share of overall PTR variation in developing countries as shown in chapter 3 (fact 3). I assume that school remoteness as measured by population density, nighttime luminosity and travel time to the closest city in chapter 3, is a good proxy of teacher costs and set $w_i/w_j = 1$.

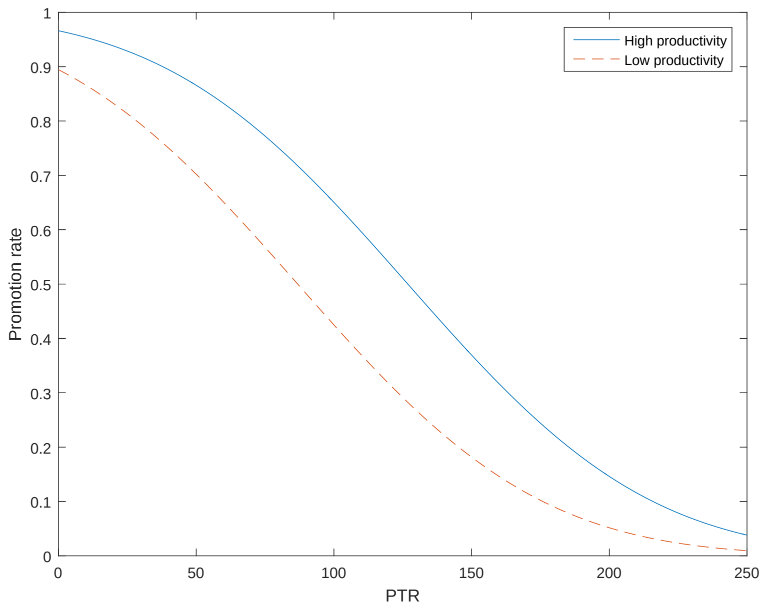


Figure 4.1: Illustration of calibrated policy function
The policy function of the low-productivity school corresponds to the case of a school with a current completion rate of 75% at a PTR of 40. The high-productivity school achieves the same completion rate at a PTR of 80.

I also consider two alternative proxies of teacher costs: actual hardship allowance payments and teacher retention rates. I find that also these can only explain a small share of PTR variation. Figure 4.2 shows the distribution of PTRs by hardship allowance category in Zambia. Teachers working at schools categorized as rural receive a rural hardship allowance that amounts to 20% of their basic salary. At so-called remote schools the allowance is 25%. The figure illustrates that there is substantial variation within each hardship allowance category and a regression of PTR on hardship allowance category confirms that only 9.8% (adjusted R^2) of PTR variation can be explained by the allowance category[3].

[3] Results for Mozambique are similar and available in the appendix. See figure D.1. Data on hardship allowance payments for additional countries has not been obtained.

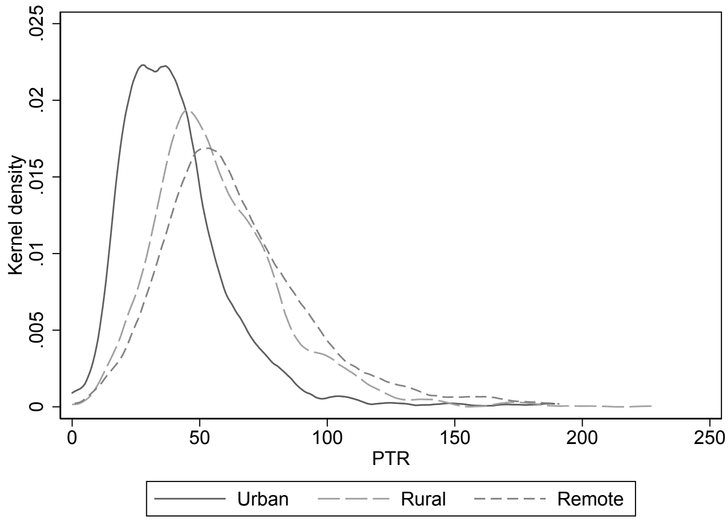


Figure 4.2: Distribution of PTRs by hardship allowance category in Zambia
Data sources: Education Management Information System, Ministry of General Education, Zambia (2014) and government payroll system, Public Service Management Division, Zambia (2014).

The second alternative proxy, teacher retention rates, explains even less of the overall PTR variation in Zambia as shown in figure 4.3. It can only explain 1.3% (adjusted R^2)[4].

[4]Results for Uganda are similar and available in the appendix. See figure D.2. Data on teacher retention rates for additional countries has not been obtained.

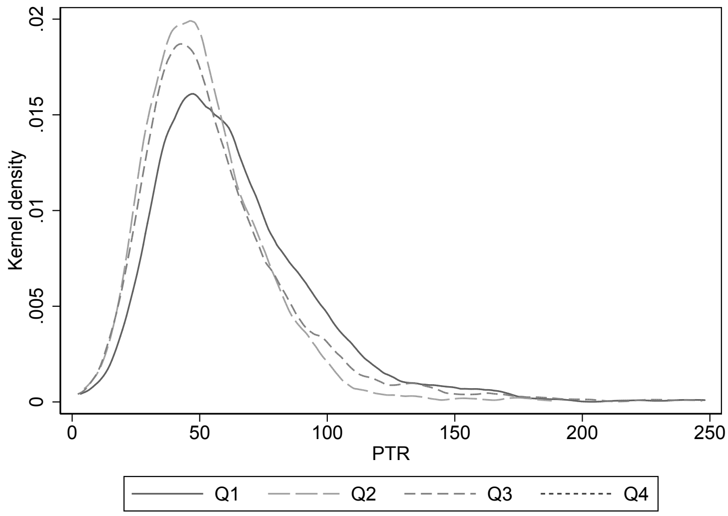


Figure 4.3: Distribution of PTRs by quartile of teacher retention rates in Zambia
Data source: Education Management Information System, Ministry of General Education, Zambia (2015).

4.2 Simulations

4.2.1 Model for simulations

The model resulting from the calibration described in the previous section can be summarized as follows:

$$\max_{T_s} \sum_s \frac{P_s}{\sum_j P_j} H_s(T_s)$$

subject to

$$\begin{aligned} H_s = 1 - \Phi\left(\Phi^{-1}(1 - c_s) + \beta\left(r_s - \frac{P_s}{T_s}\right)\right) \\ \sum_s T_s = \bar{T} \\ T_s \geq 1 \ \forall s \end{aligned}$$

This is the model used for simulation. The notation is the same as in chapter 2. $\bar{T}$ stands for the total number of teachers. Notice that the model requires that every school is allocated at least one teacher.

4.2.2 Data for simulations

To simulate the effect of alternative teacher distributions on grade promotion in public primary schools, I construct measures of grade promotion for 20 countries. I do not observe grade promotion itself in all countries. In some countries, I only observe the grade repetition rate and in others only the pass rate at a national primary school exam. In the former case, I assume dropout is zero and compute grade promotion as one minus the grade repetition rate. In the latter case, I assume that the grade promotion rate equals the exam pass rate. Table 4.1 gives an overview of the data.

Table 4.1: Measure of grade promotion by country

Country	**Measure of grade promotion**	**Missing data**
Argentina	Grade promotion rate	<1%
Cambodia	Grade promotion rate	6%
Cape Verde	1-grade repetition rate	7%
Chile	Grade promotion rate	<1%
Colombia	Grade promotion rate	14%
Djibouti	1-grade repetition rate	0%
El Salvador	Grade promotion rate	0%
Guinea-Bissau	Grade promotion rate	0%
Honduras	1-grade repetition rate	0%
India	National lower PSLE pass rate	34%
Laos	1-grade repetition rate	0%
Malawi	Grade promotion rate	<1%
Mozambique	Grade promotion rate	3%
Peru	Grade promotion rate	0%
Saint Lucia	1-grade repetition rate	0%

Table 4.1: Measure of grade promotion by country

Country	Measure of grade promotion	Missing data
Saint Vincent	1-grade repetition rate	0%
Sweden	National grade 6 examination pass rate	50%
Tanzania	National PSLE pass rate	4%
UK (England)	National grade 6 examination pass rate	14%
Zambia	Grade promotion rate	<1%

This table lists the measure of grade promotion used in each of the 20 simulation countries. PSLE stands for Primary School Leaving Examination. The third column (missing data) indicates the share of public primary schools for which the measure was not available. Reasons for missing data are country specific.

4.2.3 Smallest achievable maximum PTR rule

In many countries the allocation of teachers to schools is rule-based. These rules can be internal guidelines of Ministries of Education or formal laws. Typically they set a maximum school-level PTR that cannot be exceeded at any school (e.g. Right to Education Act in India) or a maximum class size (e.g. Malmonides' rule in Israel).

In this subsection, I simulate the distribution of teachers under the smallest achievable maximum PTR rule in each country. First, I compute the smallest threshold x that can be satisfied given the current stock of teachers and the distribution of enrolment across schools such that the PTR does not exceed x at any school. Then, I distribute teachers based on this allocation rule. Finally, I estimate the aggregate promotion gains from this counterfactual distribution using the policy function displayed in subsection 4.2.1.

Figure 4.4 illustrates the PTR as a function of enrolment under two distinct maximum PTR rules. In the first case, the overall stock of teachers is relatively large and the government can afford a maximum PTR of 25. If such a rule is strictly implemented, PTRs will at most vary between 1 and 25, and variation is likely to be smaller unless some schools have very low enollments. In the second case, the overall stock of teachers is smaller and the government can only afford a maximum PTR of 50. As the figure shows, PTRs will now vary between 1 and 50, and ceteris

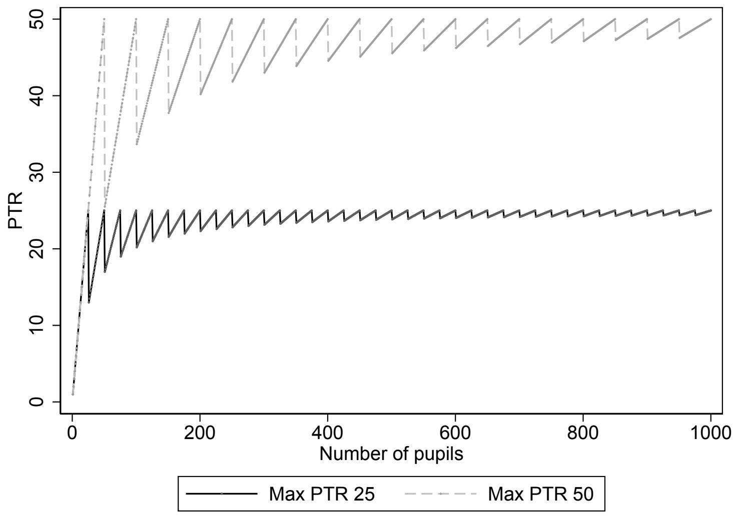
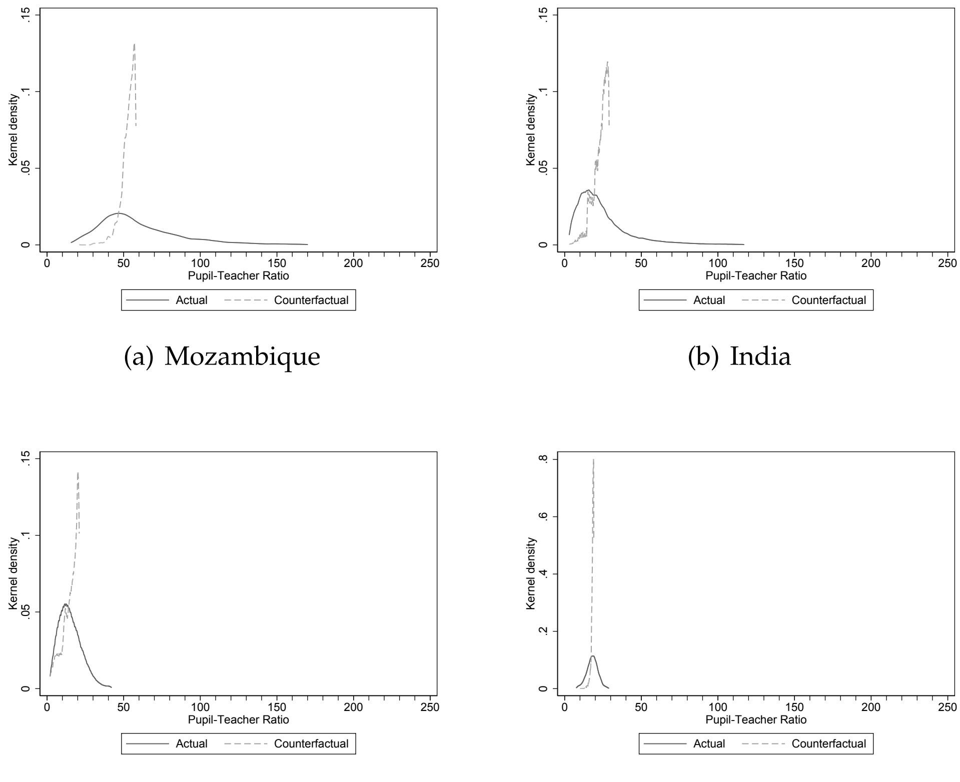


Figure 4.4: Maximum PTR rule under small and large teacher stock

paribus PTR variation will be larger than in the first case. It is evident that PTR variation mechanically depends on the overall stock of teachers and the school size (enrolment) distribution when a maximum PTR rule is used to allocate teachers. In this case, PTR variation is larger in countries with higher aggregate PTRs and smaller schools.

Figure 4.5 shows the actual and the counterfactual PTR distribution in the four example countries. It can be observed that counterfactual PTR variation is substantially smaller than actual variation in Mozambique and India, but relatively close to actual variation in Peru and the UK. At the same time, counterfactual variation is larger in the former two than in the latter two countries. This pattern is borne out across all sample countries as illustrated by the actual and counterfactual PTR standard deviations displayed in figure 4.6.

Figure 4.7 shows the percentage point gains in grade promotion for all 20 countries in the simulation sample. Countries are sorted along the x-axis by per capita income. While gains are small for countries with higher incomes, they are much larger for many low- and lower-middle-income



(a) Mozambique

(b) India

(c) Peru

(d) UK (England)

Figure 4.5: Actual and counterfactual PTR distribution under smallest achievable maximum PTR rule

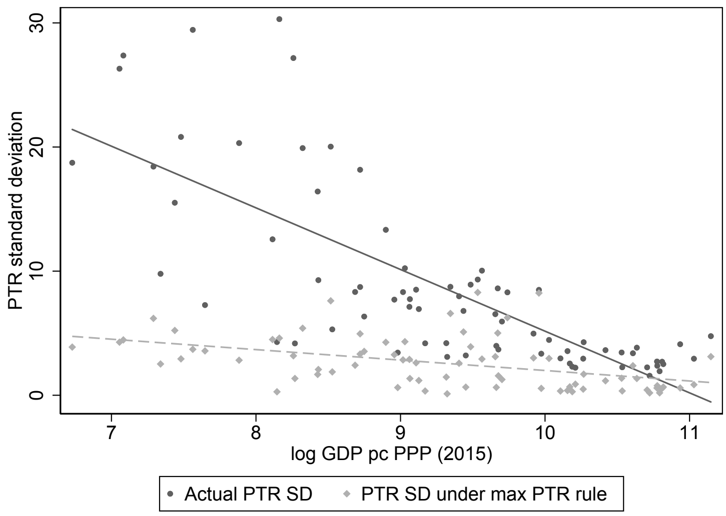


Figure 4.6: PTR variation under smallest achievable maximum PTR rule
The scatter plot contains two points for each sample country, one indicating the actual PTR standard deviation and one the counterfactual standard deviation under the smallest achievable maximum PTR rule.

countries. In India they are close to 4pp and also in Cambodia and the Sub-Saharan African countries in the sample, they are at least 1pp.

To benchmark these magnitudes I carry out two exercises. First, I ask by how much the teacher workforce would have to be increased to achieve equivalent gains if relative PTRs between schools were fixed. Table 4.2 shows how large these increases would have to be for all the countries with significant promotion gains. They are substantial in all cases. So are the associated costs that vary between 1% and 6% of total annual government education expenditure.

Second, I ask how large the gains in educational attainment are that are implied by the promotion gains. Any pupil not promoted to the next grade can either repeat the grade or drop out. Using the Young Lives survey[5], I compute the correlation between grade repetition in primary

[5]The data used come from Young Lives, a 15-year study of the changing nature of childhood poverty in Ethiopia, India, Peru and Vietnam (www.younglives.org.uk). Young Lives is funded by UK aid from the Department for International Development

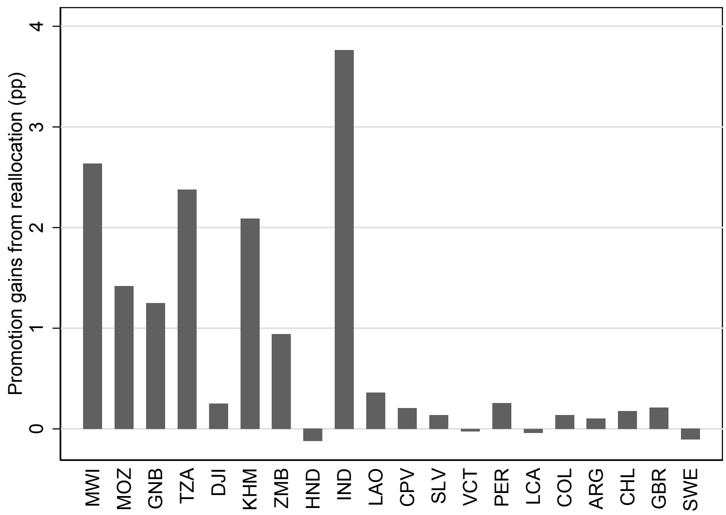


Figure 4.7: Promotion gains under smallest achievable maximum PTR rule
Countries are sorted from left to right by GDP per capita.

Table 4.2: Cost of equivalent teacher workforce increases

Country	**ΔPromotion**	**ΔWorkforce**	**Costs (USD)**	**Costs (Gvt Edu Exp)**
Cambodia	2.09pp	15.73%	43.5m/year	4.21%
Guinea-Bissau	1.25pp	8.64%	3.2m/year	5.58%
India	3.76pp	40.45%	18,170m/year	5.90%
Malawi	2.64pp	7.63%	33.8m/year	3.49%
Mozambique	1.42pp	5.92%	28.9m/year	1.33%
Tanzania	2.38pp	12.40%	191.5m/year	3.96%
Zambia	0.94pp	10.16%	139.4m/year	4.49%

ΔPromotion indicates the gain in aggregate promotion induced by teacher reallocation according to the smallest achievable maximum PTR rule. ΔWorkforce indicates the necessary increase in the teacher workforce to achieve an equivalent gain if relative PTRs between schools were fixed. Columns 4 and 5 provide the costs associated with these increases in US dollars per year and as a share of the government education expenditure. This is calculated using teacher wage data from various sources listed in table D.1 and government education expenditure data from the UNESCO Institute for Statistics and World Bank International Comparison Program Database.

school and dropout in primary school and years of education at age 22 in Ethiopia, India and Peru. Table 4.3 shows the results.

Table 4.3: Correlation with educational attainment

Country	**Grade repetition**	**Dropout**
Ethiopia	0.7	6.5
India	0.9	7.9
Peru	2.7	9.0

I divide pupils who are not promoted into repeaters and dropouts based on the actual shares in each country and consider two scenarios. For the first one I use the correlations estimated in Ethiopia, for the second the ones estimated in India. Table 4.4 shows the gains in educational attainment associated with promotion gains for all the countries with significant promotion gains in the first scenario. Table 4.5 shows them for the second scenario. The results suggest that reallocation would lead to an additional year of education for 1%-6% of children in the considered countries[6].

Table 4.4: Associated educational attainment gains - Scenario 1

Country	**ΔPromotion**	**ΔYrs Educ**	**ΔYrs/Child**
Cambodia	2.09pp	143.9K	0.046
Guinea-Bissau	1.25pp	9.0K	0.020
India	3.76pp	9189.6K	0.036
Malawi	2.37pp	64.3K	0.014
Mozambique	1.42pp	255.1K	0.033
Tanzania	2.38pp	754.2K	0.051
Zambia	0.94pp	43.9K	0.010

(DFID). The views expressed here are those of the author. They are not necessarily those of Young Lives, the University of Oxford, DFID or other funders.

[6]Data on child population between the age of 4 and 15 is from the UNESCO Institute for Statistics and World Bank International Comparison Program Database

Table 4.5: Associated educational attainment gains - Scenario 2

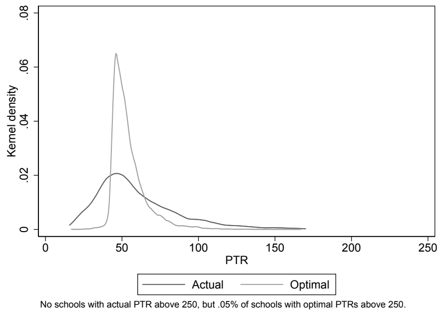
Country	**ΔPromotion**	**ΔYrs Educ**	**ΔYrs/Child**
Cambodia	2.09pp	175.6K	0.056
Guinea-Bissau	1.25pp	11.0K	0.024
India	3.76pp	11216.2K	0.044
Malawi	2.37pp	89.0K	0.018
Mozambique	1.42pp	312.2K	0.040
Tanzania	2.38pp	920.3K	0.062
Zambia	0.94pp	54.4K	0.012

4.2.4 Optimal allocation

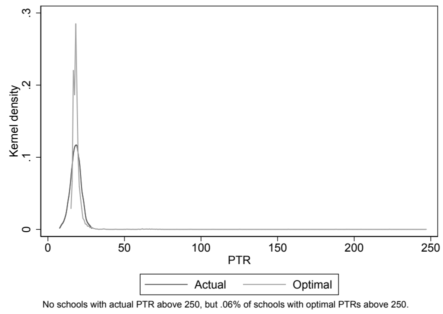
In this subsection, I ask how large gains from allocating teachers optimally would be. I numerically solve the maximization problem displayed in section 4.2.1 using a combination of direct search and gradient-based tools. For computational reasons, I allow for divisibility of teachers[7]. Figure 4.8 shows the actual and the optimal distribution of PTRs in England and Mozambique. In the optimum, very unproductive schools obtain only the minimum of one teacher. This leads to some schools with extremely high PTRs in both countries. In addition to this, it stands out that the optimal allocation implies less dispersion than is actually observed. However, in Mozambique the optimal dispersion is significantly larger than in the UK.

Figure 4.9 plots the gains from teacher reallocation. Gains from implementing the optimal allocation are substantially larger than gains from the simple maximum PTR rule discussed previously. Once again, they are large in low- and lower-middle-income countries whereas they are comparatively small in upper-middle- and high-income countries.

[7]Solving this high-dimensional optimization problem is computationally much more demanding when adding an integer constraint for each school.



(a) Mozambique



(b) UK (England)

Figure 4.8: Actual and optimal PTR distribution

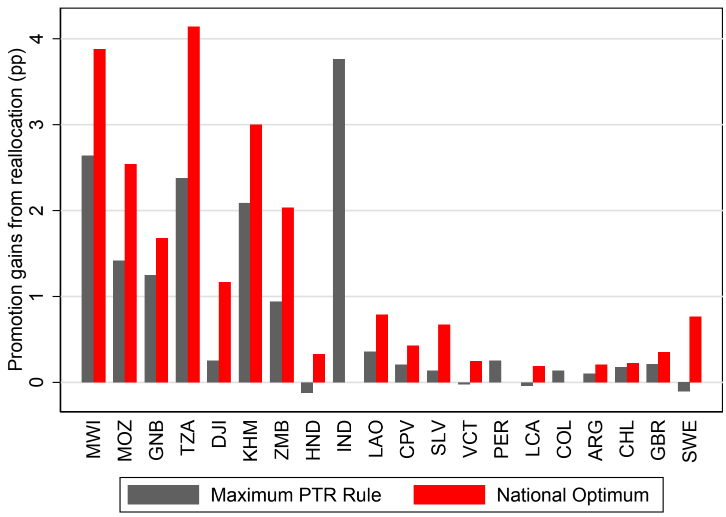
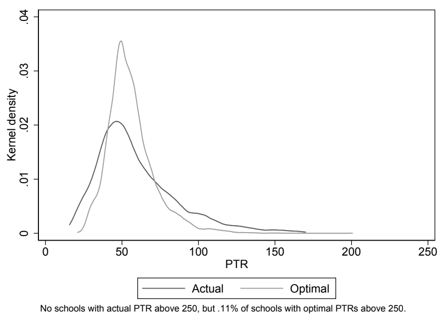


Figure 4.9: Promotion gains under optimal allocation
Countries are sorted from left to right by GDP per capita. Due to computational constraints results for Colombia, India and Peru are not available.

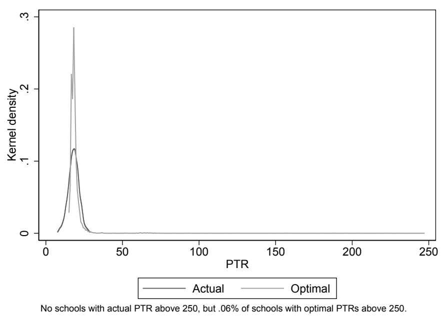
4.2.5 Optimal allocation within subnational units

The final counterfactual simulation only allows for teacher reallocation within subnational units and asks how large promotion gains from switching to an optimal allocation within these would be. In practice, improving teacher allocation within smaller units may be more easily feasible than a nationwide initiative. Given the large PTR variation within subregions documented in chapter 3, such efforts are likely to go a long way towards improving educational outcomes. For the purpose of this simulation, I define subnational units such that geographic size is as comparable as possible across countries. This means I use second-tier administrative units in large countries such as India and Mozambique, first-tier administrative units in mid-size countries such as Cambodia or UK, and the entire country for small island nations such as Cape Verde or Saint Lucia[8].

Figure 4.10 shows the actual and the counterfactual distribution of PTRs in Mozambique and the UK. The counterfactual dispersion is smaller than the actual one, but only marginally so in England. Moreover, in both countries it is larger than under the two previous counterfactuals.



(a) Mozambique



(b) UK (England)

Figure 4.10: Actual and counterfactual PTR distribution under optimal allocation within subnational units

[8]See table D.2 for details.

The gains from reallocation are shown in figure 4.11. In all countries, they are larger than gains from implementing the smallest achievable maximum PTR rule. They are also somewhat smaller than gains under the nationwide optimal allocation, but not by very much in most countries, indeed suggesting optimal allocation within subnational units could have large effects on aggregate education outcomes.

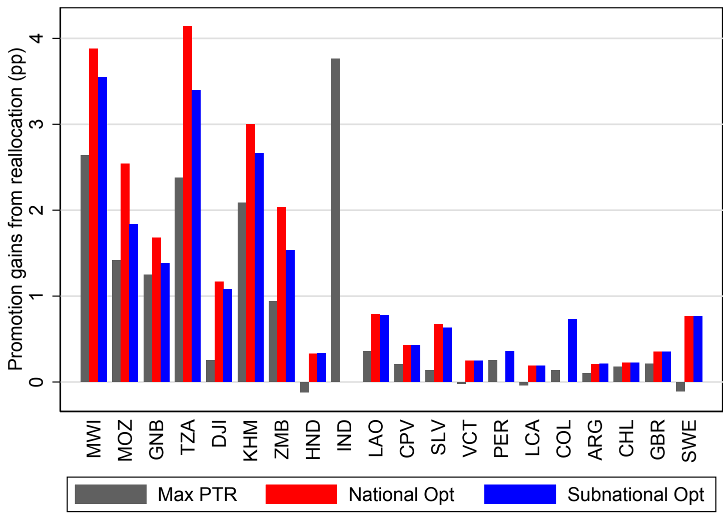


Figure 4.11: Promotion gains under optimal allocation within subnational units
Countries are sorted from left to right by GDP per capita. Due to computational constraints results for India are not available.

4.3 Robustness

All three counterfactual simulations suggest that gains from teacher reallocation in developing countries could be substantial. In this section, I briefly examine the sensitivity of the presented results.

The magnitude of the estimated effects directly depends on the choice of the parameter β. Gains are overestimated if the effect of PTR on grade promotion is smaller than assumed. Figure 4.12 shows promotion gains under the smallest achievable maximum PTR rule with β half as big as previously assumed. While gains are smaller, they are still significant in the same countries as before. So, the conclusion that teachers are misallocated in developing countries in South Asia and Sub-Saharan Africa does not hinge on the precise magnitude of the parameter β. Indeed, figure 4.13 illustrates that gains would be substantially larger if β was twice as large as previously assumed.

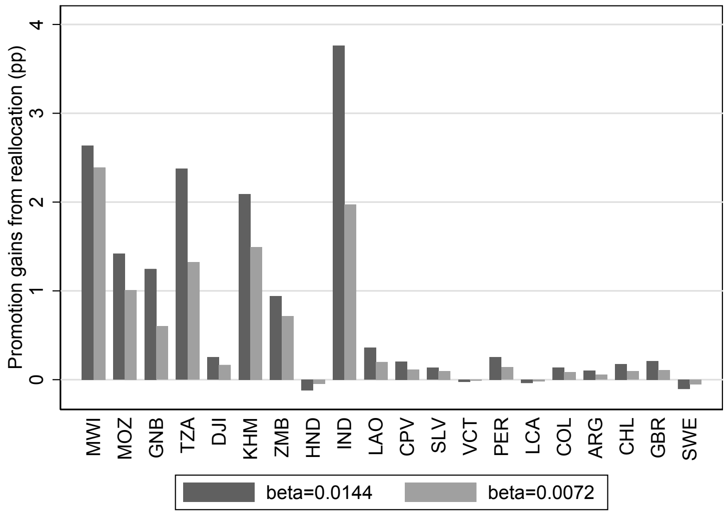


Figure 4.12: Promotion gains under smallest achievable maximum PTR rule - Small β
Countries are sorted from left to right by GDP per capita.

Gains are also overestimated if teacher costs are not equal across schools, but instead positively correlated with PTRs. In this case, the total number of teachers that could be employed under the smallest achievable maximum PTR rule would be smaller than the current stock of teachers (given a fixed budget for teacher compensation). Hence, the smallest achievable maximum PTR would be larger than the one computed in section 4.2.3, and aggregate promotion gains therefore smaller.

Finally, the effect of PTR on promotion may depend on school characteristics and it is a priori ambiguous how this would affect the estimated gains. For example, the marginal benefit from an additional teacher at a well equipped school could either be larger or smaller than

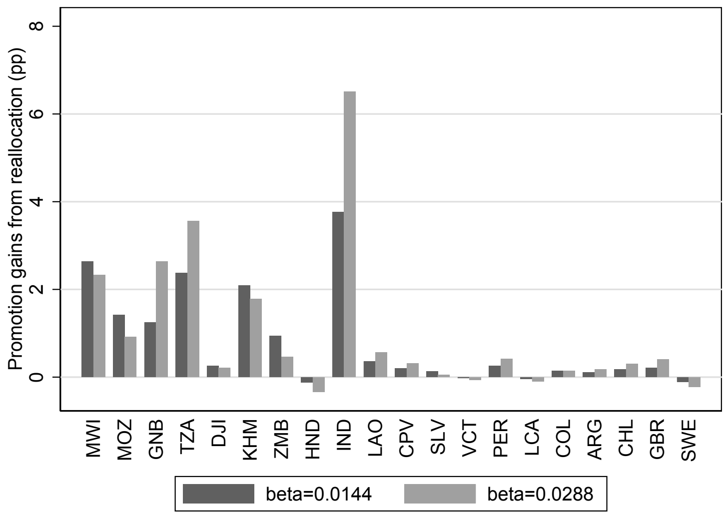


Figure 4.13: Promotion gains under smallest achievable maximum PTR rule - Large β
Countries are sorted from left to right by GDP per capita.

at a poorly equipped school. Muralidharan and Sundaraman (2013) show that the treatment effect of their RCT is heterogeneous by school infrastructure[9] and proximity to facilities[10][11]. They find that effects are larger for more remote and less well-equipped schools. I construct the school infrastructure measure used by the authors for all public primary schools in India using data from the Indian District Information System for Education (2015) and show that school infrastructure is negatively

[9]Infrastructure is measured by an index summing six dummy variables that indicate the existence of a brick building, a playground, a compound wall, a functioning source of water, a functional toilet, and functioning electricity.

[10]Proximity to facilities is measured by an index summing eight variables (each coded from 1-3) indicating proximity to a paved road, a bus stop, a public health clinic, a private health clinic, public telephone, bank, post office, and the mandal educational resource center. A higher value of the index indicates being further away from these facilities.

[11]Notice that this is in line with misallocation of teachers across schools. If the effect of adding a teacher was equal across schools, the marginal product would be equalized and the current allocation could be optimal.

correlated with PTRs[12]. This suggests that reallocation gains from implementing the smallest achievable maximum PTR rule in India would be even larger than estimated above because gains from adding teachers to high-PTR schools would be larger than losses from withdrawing teacher from low-PTR schools.

[12] A reduction in the infrastructure index by one is associated with a PTR increase by 1.9.

Chapter 5

Causes: A case study from Zambia

The distribution of teachers across public primary schools in Zambia is representative of that in many developing countries. As figure 5.1 shows, pupil-teacher ratios (PTRs) vary widely across public primary schools. While the national aggregate PTR is 44.2, the bottom 10% of schools have PTRs below 29.9 and the top 10% have PTRs above 101. Approximately 475,000 pupils, about 16% of the public primary school population, attend schools with a PTR above 80. A large share of the variation in PTRs is within districts. A decomposition of the PTR variance reveals that the within-district variation (30.5) is far larger than the cross-district variation (12.1) .

This chapter constitutes an investigation of the causes of PTR variation in Zambia. It traces the disparities in staffing levels back to a set of interlinked administrative issues. These include:

1. An ineffective teacher allocation policy

2. Non-compliant deployment and transfers

3. Payroll mismatch

4. Weaknesses in the budgeting process for teacher positions

These issues are dealt with in detail in following sections.

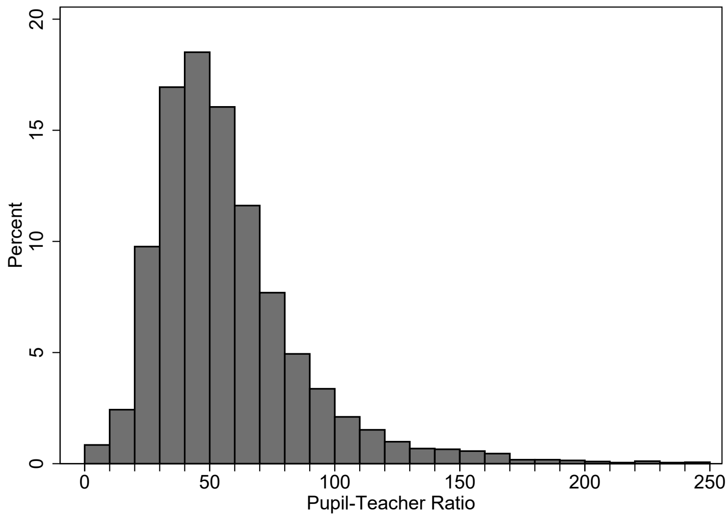


Figure 5.1: Distribution of PTRs across public primary schools in Zambia
DataSource: Education Management Information System (2017), Ministry of General Education, Zambia.

5.1 Ineffective teacher allocation policy

The Standards and Evaluations Guidelines of the Zambian Ministry of General Education posit that no school should have a PTR greater than 40. However, this rule is largely not followed, and 73% of public primary schools have PTRs greater than the required maximum. At the same time, 21% of schools have more teachers than the minimum number required to meet this rule. Figure 5.2 shows schools' current staffing levels versus the minimum number of teachers that they would need to meet the Ministry's maximum PTR rule. As it indicates, while many schools have less than the minimum prescribed number of teachers (below the 45° line), there are also many that have well over this minimum (above the 45° line). Many of these schools could have teachers transferred to schools with fewer teachers than necessary and still have a PTR in line with the government directive. However, even if excess teachers from these schools were reallocated to schools in need of teachers, approximately 12,500 new primary school teachers would have to be hired in addition to the existing stock of teachers in order to be able to meet the target of a maximum PTR of 40 for all public primary schools.

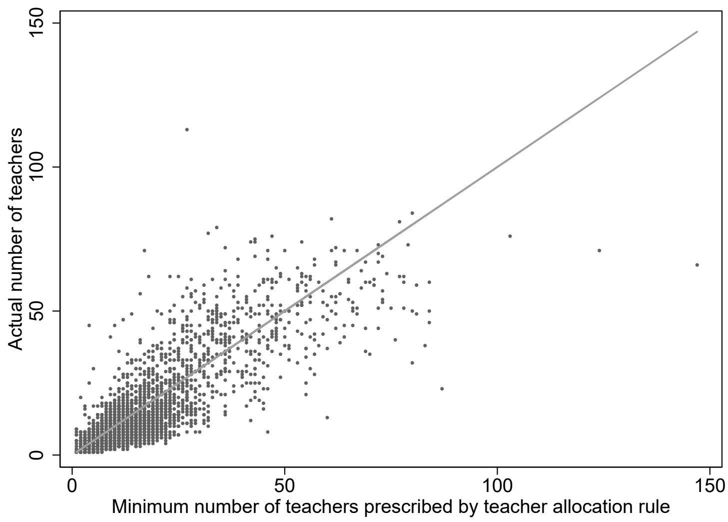


Figure 5.2: Actual versus prescribed number of teachers across public primary schools
Data source: Education Management Information System (2017), Ministry of General Education, Zambia. The grey line is a 45Â° line.

5.2 Non-compliant deployment and transfers

The deployment of teachers does not appear to be particularly responsive to current staffing needs. One would expect that more teachers are deployed to areas with higher PTRs. But figure 5.3 shows that there is very little correlation between the number of teachers a school would need to achieve a PTR in line with the government's mandate of 40 in 2013 and the number of new teachers deployed to that school in 2014. As the figure indicates, most schools do not receive any new teachers. However, the schools that do receive teachers are often not those in the most need. In fact, many schools that already have more teachers than necessary to achieve the PTR rule receive new teachers in deployment, rather than those teachers being sent to understaffed schools. These schools, as well as those that previously had PTRs above 40 but receive more teachers than necessary to achieve the PTR rule, are represented by red crosses in the plot.

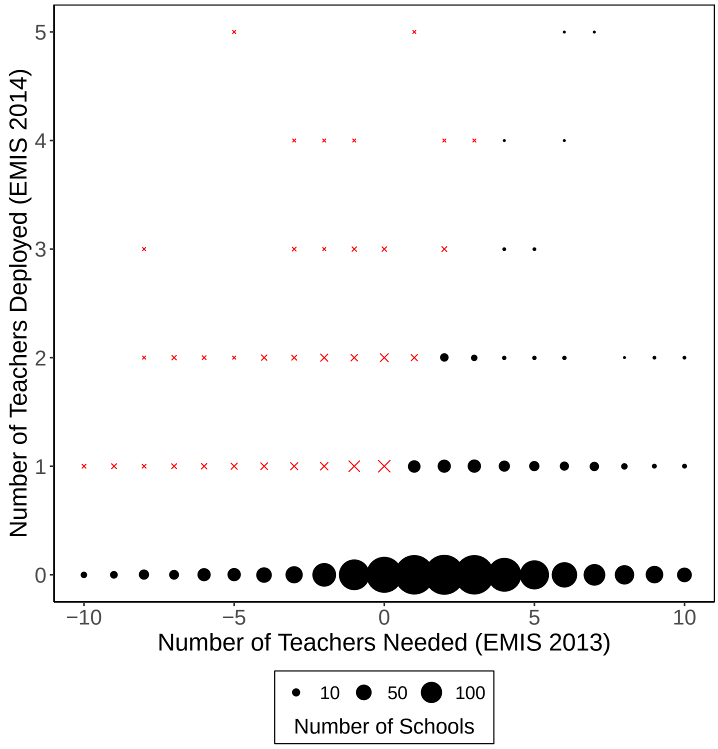


Figure 5.3: Need for teachers and deployment of teachers in the following year
Marker size indicates number of schools. Red crosses indicate cases where more teachers are deployed than necessary to achieve a PTR of 40. Data source: Source: Education Management Information System (2013-2014), Ministry of General Education, Zambia.

The teacher transfer process can also contribute to unbalanced staffing patterns if transfers go from relatively understaffed to relatively overstaffed schools. Analysis of teacher movement in EMIS between 2010 and 2016 indicates that while the majority (approximately 60%) of transfers places a teacher into a school with a higher PTR than the one they come from, a large share of transfers (approximately 40%) move teachers into schools with lower PTRs than those they come from. Additionally, approximately the same number of transfers occurs from schools with

PTRs above 40 to schools with PTRs below 40 (19%) as the inverse (20%). This suggests that overall transfers hardly contribute to equalizing PTRs. Moreover, it raises the question why so many transfers from understaffed schools to well-staffed schools occur. In principle, there are regulations that limit the number of transfers and their impact on PTRs. Transfer requests must be approved by senior officers, and the government has imposed a minimum holding period at a new school before a teacher can transfer. This holding period has recently been increased from two to four years. However, even with the shorter minimum, it frequently has not been respected. As figure 5.4 indicates, over half of the teachers that transferred at some point between 2010 and 2016 did so before two years had elapsed, and nearly 90% did so before four years . Anecdotal evidence suggests a number of causes for this including health issues that are sometimes exaggerated, teachers obtaining transfers to be with spouses, and teachers using social and political connections to obtain transfers to preferred locations.

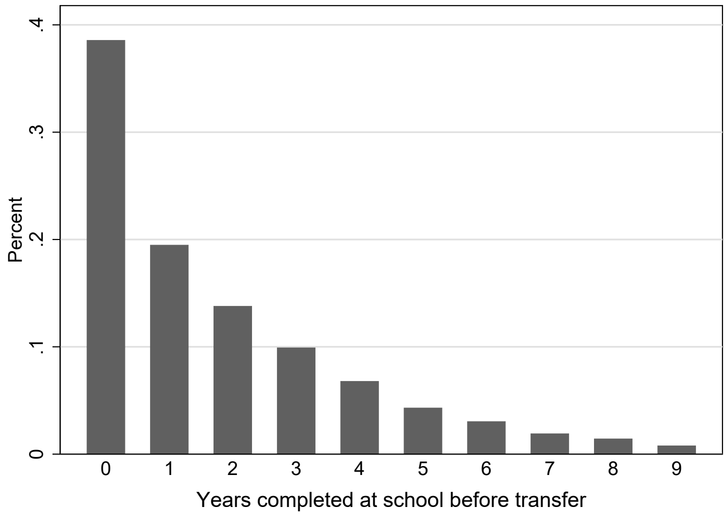


Figure 5.4: Distribution of holding periods (completed years) among teachers who transferred between 2010 and 2016
Data source: Source: Education Management Information System (2010-2016), Ministry of General Education, Zambia.

5.3 Payroll mismatch

Transfers that are not in compliance with official guidelines are also likely to contribute to significant levels of payroll mismatch, another major administrative obstacle to teacher allocation. Payroll mismatch occurs when staff does not work at the organizational unit they are listed at in the government payroll system (their paypoint). Quantifying the exact magnitude of payroll mismatch in the education sector is difficult, however several studies have provided similar estimates. In a 2014 report on a sample of 88 schools in four provinces, the Office of the Auditor General (OAG) found that up to 60% of teachers do not work at the location where they are paid. A 2016 survey of 158 rural schools I conducted myself placed this number at 40%. Analysis of 2017 staff returns from two provinces in which district officials indicate the working location and paypoint of each teacher reveals a payroll mismatch range between 43% and 77% across districts.

Payroll mismatch is a major problem with regards to the allocation of teachers because a school may be understaffed while payroll does not show any vacancies for the school. In this case, some of the teachers on the school's payroll presumably work at other schools (or not at all). Figure 5.5 illustrates this problem. It plots the number of teachers reported in the payroll system versus the actual number of teachers as reported in the Education Management Information System (EMIS) in 2014 for each public primary school. If there was no payroll mismatch, the points would all lie on the grey 45° line. But as the graph shows, many schools (30%) have more teachers in EMIS than on their payroll (indicating overstaffing relative to payroll), and an even larger number (61%) have more teachers on payroll than in EMIS (indicating understaffing relative to payroll). Most teachers working at schools where EMIS teacher counts exceed the number of paypoints are likely occupying paypoints at understaffed schools, thus impeding adequate deployment of teachers to these schools.

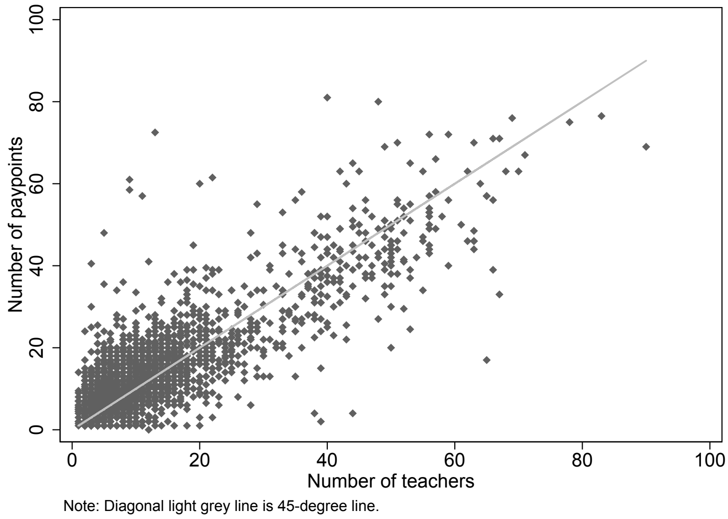


Figure 5.5: Number of paypoints and teachers by school
Data sources: Education Management Information System (2014), Ministry of General Education, Zambia, and government payroll system (2014), Public Service Management Division, Zambia. Figure based on 5,293 schools that could be matched between EMIS and the payroll system by school name, 87% of schools in the payroll system.

5.4 Weaknesses in the budgeting process

The prevalence of payroll mismatch means that there are substantial discrepancies between teachers' assigned placements and their actual working locations. However, this does not mean that eliminating payroll mismatch would eliminate the observed variation in staffing levels across schools. In fact, even if actual staffing levels perfectly followed payroll, PTRs would still vary substantially across schools. Figure 5.6 illustrates this by comparing the distribution of actual PTRs across public primary schools to the distribution of pupil-teacher paypoint ratios (henceforth sanctioned PTRs) across the same schools. The figure shows that there is not only significant variation in actual PTRs, but also in sanctioned PTRs. 40% of schools have sanctioned PTRs above 40 and 60% of schools have sanctioned PTRs below 40.

One of the main factors behind the dispersion in sanctioned PTRs appears to be that establishment registers determining the human re-

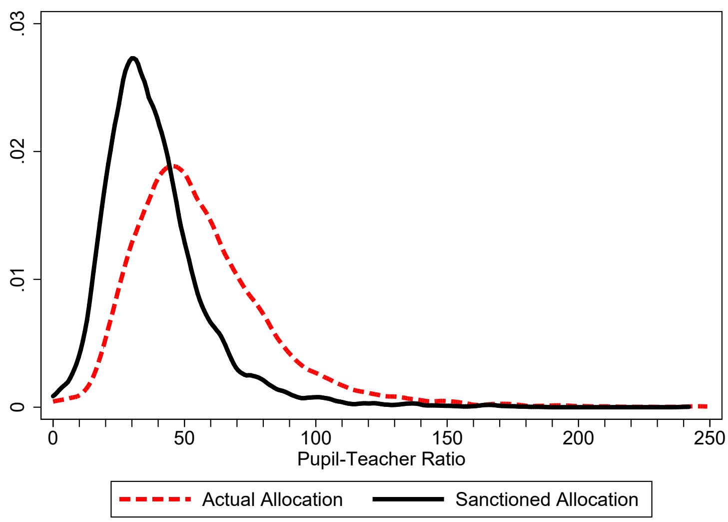


Figure 5.6: Actual and sanctioned distribution of PTRs across public primary schools in 2014
Data sources: Education Management Information System (2014), Ministry of General Education, Zambia, and government payroll system (2014), Public Service Management Division, Zambia. Actual PTRs are derived from EMIS 2014 and sanctioned PTRs are computed based on 2014 payroll data. Sanctioned PTRs are defined as pupil-teacher paypoint ratios.

source budget of each school position by position are rarely updated. Establishment registers determine the human resource budget of each school position by position. However, once a school has been opened, its establishment is rarely adjusted to reflect changes in enrollment over time. Between 2012 and 2016, only 10% of schools had any update, and only 8% had an update that added teachers. It is also unclear how closely establishment updates reflect changes in enrolment. While schools that gained teachers did have larger increases in enrolment than schools without a change in their establishment, 90% of schools with an increase in enrolment did not see an increase in their establishment and 47% of schools that gained teaching positions saw decreases in enrolment.

Additionally, newly opened schools can take a long time to receive an establishment. Schools are frequently opened before funds are officially allocated towards fully staffing it. Even when funds are available, the process to approve an establishment is complex. Approval is needed at multiple levels of the Ministry of Education, the Public Service Management Division, and the Ministry of Finance, which opens many opportunities for the process to stall. It is difficult to estimate how many schools are functioning without an establishment, but collaborative fieldwork with the Human Resource Department of the Ministry of Education and the Zambia Education Sector Support Technical Assistance (ZESSTA) in one district of Zambia found that 38% of schools did not have an establishment. Notice that figure 5.6 does not take any facilities without establishment registers into account as they are not observed in payroll. Therefore, the true dispersion in sanctioned PTRs is even larger than shown in the figure.

Missing and outdated establishment registers are not only a problem for teacher allocation in and of themselves, but they also cause payroll mismatch and thus additionally affect teacher allocation indirectly through this channel. This is because district education offices need to send teachers from schools with establishment registers to schools that do not have sufficient, or any, paypoints in order to guarantee the operation of these schools.

5.5 Discussion

In chapter 3, it was shown that school remoteness can only explain a small share of overall variation in PTRs across public primary schools in developing countries. This chapter suggests that instead lack of managerial capacity and weak enforcement of existing policies are key drivers of PTR variation, at least in the case of Zambia.

Research and government reports from other developing countries paint a similar picture. For example, Diompy (2014) describes how favoritism plays an important role in teacher allocation decisions in Senegal and criticizes the lack of clear transfer policies. Cummings and Tahirou (2016) find that in Niger, "there is a lack of regulations, accountability mechanisms and sanctions governing the distribution of teachers". Ramachandran et al. (2018) describe a patronage-based teacher allocation system in India where politicians and bureaucrats transfer teachers for political motives regardless of school need. Similarly, Asim et al. (2017) document how in Malawi ”teachers leverage informal networks and political patronage to resist placement in remote schools, while administrative officials are unable to stand up to these formal and informal pressures, in part because of a lack of reliable databases and objective criteria for the allocation of teachers”. Weaknesses in allocation management are also reported by governments themselves. For example, Cameroon’s strategic education sector plan states that ”imprecise management, in particular with regards to the allocation of teachers across schools, results in both an efficiency and an equity problem” (Ministere de l’economie, de la planification, et de lâamenagement du territoire, Republique du Cameroun 2013). A recent strategic document from the government of Bangladesh announces a ”major shift (...) to a demand or need-based deployment of resources, including teachers” and emphasizes that ”clear criteria will be applied to determine the actual need for new teachers on a school-by-school basis” (Ministry of Primary and Mass Education, Government of the People’s Republic of Bangladesh 2015).

The described causes of PTR variation are unlikely to be confined to the education sector. In developing countries, lack of managerial capacity and enforcement problems are likely to be present across different ministries. Hence, the uncovered human resource allocation problem is likely to extend to other public sectors, such as health, law enforcement and administration. The next chapter takes a look at the distribution of health workers.

Chapter 6

Human resource misallocation in other public sectors? Evidence from the staffing of Zambian and English primary care facilities

Many developing countries suffer from a lack of health resources. Health staff shortages are an area of particular concern. The WHO recommends a minimum of 4.45 doctors, nurses, and midwives per 1000 population (Scheffler et al. 2018). Figure 6.1 shows aggregate staffing levels per 1000 population relative to this benchmark across countries by per capita income. The overall lack of health workers in low- and middle-income countries is appalling. Nonetheless, some have claimed that ”imbalances of health workers are a bigger issue than national deficiencies” (Lemiere et al. 2013) in developing countries, thereby hinting at severe inefficiencies in human resource allocation in the health sector. Empirical evidence on the distribution of health workers, however, is scarce and confined to analyses at the level of subnational administrative units[1].

This chapter sets out to compare the distribution of health workers in Zambia and England at the micro-level, i.e. the health facility level.

[1]See Appiah-Denkyira et al. (2013), Dussault and Franceschini (2006), Ferrinho et al. (2011), Lemiere et al. (2011), and Munga and Maestad (2009) for available evidence.

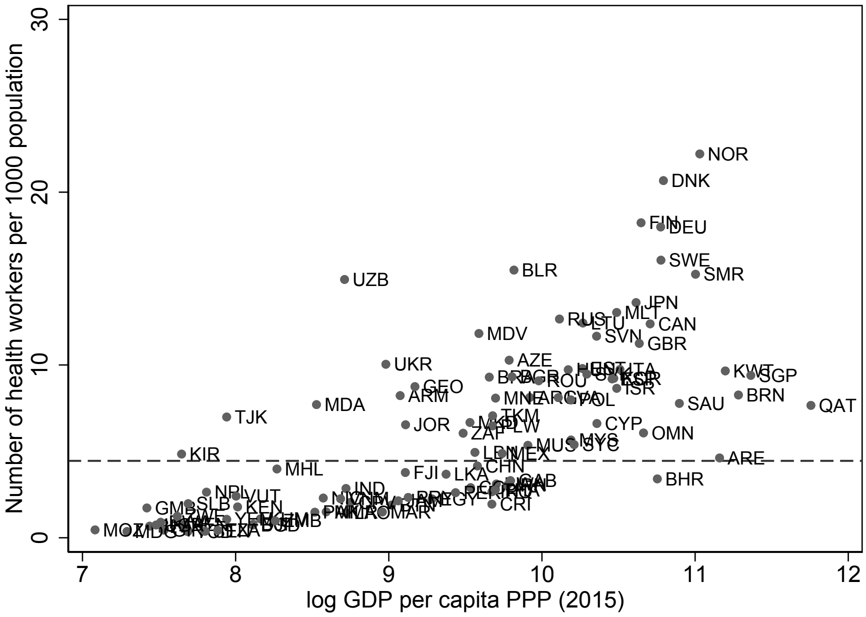


Figure 6.1: Doctors, nurses, and midwives per 1000 population across countries by income

Health worker density data is from the WHO's Global Health Workforce Statistics. Numbers are the most recent available for each country. Per capita income data is from the World Bank International Comparison Program database. The dashed horizontal line indicates the WHO recommendation of 4.45 doctors, nurses, and midwives per 1000 population.

Capturing staffing variation within subnational units is important for two reasons. First, a substantial share of cross-facility staffing variation may be within subnational units, as shown for the education sector in chapter 3. In this case, analyses that compare aggregate staffing levels between subnational units seriously underestimate staffing imbalances. Second, capturing local variation is also important because local understaffing could have large negative externalities as it may limit the capacity to contain disease outbreaks locally. The 2014 Ebola outbreak in West Africa is a recent reminder of the magnitude such externalities can take on[2].

I propose a general approach to describe the spatial distribution of health workers relative to population which is applicable across countries

[2]See Panjabi (2016) for more information.

of all income levels as long as data on facility locations and facility staffing exists. While the presented approach can be employed across all health care sectors, the focus of this chapter is on public primary care and therefore only medical staff at public primary care units is taken into account. Distributional patterns in England serve as a reference for those in Zambia.

A key challenge to documenting the relative distribution of health workers in developing countries at the micro-level is estimating catchment populations of health facilities. Corresponding data is frequently not available and when available, often of uncertain quality. I propose two estimation methods and assess their performance through comparison with official catchment population headcounts for subsets of facilities where such data is available. Both estimation methods rely on high-resolution population data from the Gridded Population of the World[3] which reports population estimates at the level of approximately one-by-one kilometer grid cells for all countries. The two methods differ in the way grid cells are assigned to facilities. Each cell is assigned to the closest facility, but closeness of facilities is firstly measured by straight-line distance and secondly by travel time. Both methods explain about one fourth of the cross-facility variation in catchment population headcounts in England and one tenth of it in Zambia. Estimates based on travel time slightly outperform those based on straight-line distance in both countries. Lower data quality in Zambia is likely to account for the lower explanatory power of both of the methods in Zambia compared to England.

Independent from the source of facility catchment population, whether taken from official headcounts or estimated as described, I find large variation in access to health workers across the population in Zambia. The distribution is marked by a long right tail of people living in areas with high population-health-worker ratios (PHRs). In contrast, in England PHRs vary little and there is hardly any right tail. Modal PHRs, however, are similar in both countries. A spatial decomposition of the

[3]See http://sedac.ciesin.columbia.edu/data/collection/gpw-v4 for details.

variation in relative staffing across facilities reveals that a large share of the variation in Zambia is within districts. Finally, regressions of PHR on measures of facility remoteness show that only a small share of overall staffing variation can be explained by remoteness. All these findings mirror facts 1 and 2 presented for the education sector in chapter 3, thus suggesting that human resource allocation could be equally inefficient in the health sector of developing countries.

The chapter proceeds as follows. Section 2 describes the data used in this chapter. In section 3, the two different methods to estimate catchment populations are discussed. Section 4 presents the results and section 5 discusses them.

6.1 Background and data

6.1.1 Zambian primary care facilities

Several administrative data sets on Zambian health facilities were combined to build a comprehensive database that includes the location and staffing for the universe of primary care facilities in the country. For this purpose primary care units were defined as health posts and health centres. These two types of facilities represent the bottom two tiers of the five-tier Zambian health system. The excluded top three tiers are comprised of hospitals of different levels of specialization.

In 2017, EQUIP Zambia conducted a census of all health facilities in the country. This census provides the most complete listing of health facilities in Zambia and contains GPS coordinates for all facilities. However, there is a small number of primary care facilities (172) that EQUIP did not manage to visit for logistical reasons. Coordinates for 94 of these could be recovered from the 2007 JICA health facility census. In the end, only 78 primary care facilities remained without coordinates and will not be taken into account below as location information is essential for the analysis. The total number of primary care facilities considered below amounts to 2490.

In addition to facility location information, the EQUIP census also collected official catchment population headcounts for a subset of primary care facilities (68%). This data is used for comparison with estimated catchment populations. Official headcounts are conducted by facility staff and are supposed to include all inhabitants in the official catchment area of a facility. Official catchment areas are typically composed of a number of settlements assigned to a facility, however it is unclear to what extent these overlap with actual catchment areas as patients are not restricted in their facility choice and settlements need not be part of any official catchment area[4].

Facility staffing data comes from the Ministry of Health's Human Resource Information System (HRIS). The data used is from January 2018. HRIS is restricted to public health workers and therefore, it does not contain information on the staffing of private facilities. However, only 4.9% of primary care facilities in Zambia are private. HRIS staffing data was matched to facility data by facility name. This way staffing data was successfully merged to 74% of primary care facilities. Unmatched public facilities are mainly due to missing facility names in HRIS in certain districts. In the following analysis, all facilities with and without matched staffing data are taken into consideration for the construction of catchment areas.

6.1.2 English primary care facilities

For the purpose of this paper and for comparability with Zambia, I define primary care units in England as General Practitioner (GP) practices[5]. In England, every resident is eligible for free primary care services through the National Health System (NHS) and needs to register with a GP practice for this purpose. At any one point in time, a resident cannot be registered at more than one GP practice and the choice of GP practices

[4]Detailed information on the assignment of settlements only exists at a decentralized level and could not be obtained for this study.

[5]Note that this definition excludes certain primary care providers such as dentists, opticians, and pharmacists.

they can register at is restricted by the location of residence. Only residents within a GP's practice area are entitled to register. This means that residents can only choose between a small number of GP practices. While it is not mandatory to register with a GP, nearly every resident is registered. In fact, in 2015 the total number of registered patients with NHS exceeded official population estimates from the Office for National Statistics[6]. In this chapter, the number of patients registered with each practice is used as the official catchment population.

Data on the location (GPS coordinates), staffing and registered patients of all GP practices in England is published by NHS Choices and data from 2014 was downloaded from the UK government's open data portal[7]. While location and staffing data are provided at the GP branch level, registered patients are only available at the GP practice level. Overall the data contains 9,847 GP branches belonging to 8,157 GP practices and location information is available for 96.7% of branches. As in the case of Zambia, facilities without location information cannot be considered in the subsequent analysis. In the following section, catchment areas are constructed at the branch level, but catchment populations are aggregated at the practice level in order to allow for comparison with the number of registered patients at each practice.

6.1.3 Complementary data

Population

The Gridded Population of the World (Version 4) from Columbia University is the primary source of high-resolution population estimates for this paper[8]. It provides population estimates within 30 arc-second grid cells (approximately $0.8km^2$ throughout Zambia and $0.5km^2$ throughout

[6]See Baker (2016) for details.

[7]See https://data.gov.uk/. Data was downloaded on 09/02/2017.

[8]WorldPop from the University of Southampton was considered as an alternative source of high-resolution population data for Zambia. As it does not cover England, however, it could not be used for comparative analysis. Results for Zambia based on WorldPop are available upon request.

England). Its main input is national census data. Population from the smallest available administrative area (constituencies in Zambia and output areas in England) is assumed to be evenly distributed within these areas, with adjustments made for borders and landforms such as bodies of water[9].

Travel time

Based on geographic and road data from Google and OpenStreetMap, the Malaria Access Project (MAP) at Oxford University has produced a data set dividing the world into grid cells of the same size as the Gridded Population of the World indicating the difficulty to travel through each grid cell[10]. This data is used to determine which facility can be reached in the shortest time from each population grid cell.

6.2 Estimation of catchment populations

I use two different approaches to estimate catchment populations and compare results to official estimates. Official estimates are based on headcounts by facility staff in Zambia and the number of registered patients in England. While the latter is available for the universe of GP practices in England, the former is only available for 68% of primary care facilities in Zambia.

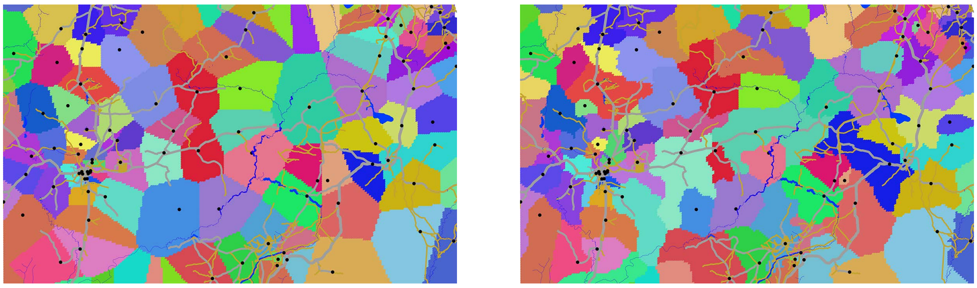
The first approach assigns each population grid cell to the closest facility based on the straight-line distance between the centre point of the grid cell and the facility. The second approach assigns grid cells to facilities based on shortest travel time as determined by the MAP data previously described[11] In both cases, the catchment population of a given facility is computed as the sum of population over all cells in

[9]See http://sedac.ciesin.columbia.edu/data/collection/gpw-v4 for details.

[10]See Weiss et al. (2018) and https://map.ox.ac.uk/research-project/accessibility_to_cities/ for details.

[11]When multiple facilities fall into the same MAP cell, catchment populations are evenly distributed between them. Results are robust to distributing catchment populations proportional to facility staffing between facilities in these cases instead.

the facility's catchment area[12]. Figure 6.2 demonstrates the similarities and differences between these two approaches in Zambia. While the straight-line approach yields fairly regular polygons (6.2(a)), travel time-based catchment areas conform more closely with natural boundaries, and extend along major roads (6.2(b)). Overall, 30% of Zambia's and the UK's area and 31% and 32% of their respective populations lie in different catchment areas when using one method instead of the other.



(a) Straight-line distance

(b) Travel time

Figure 6.2: Health facility catchment areas in Northern Zambia by construction method

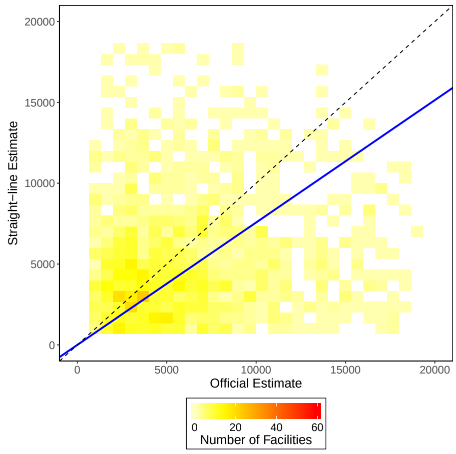
Dots indicate primary care facilities. Catchment areas represented by different underlying colors. Grey lines are roads and blue lines are bodies of water.

Figure 6.3 shows the correlation between derived catchment population estimates and official numbers for all facilities where the latter are available. Independent from the estimation method, estimated catchment populations are positively correlated with facility headcounts. In England the correlation between official and estimated catchment populations is stronger than in Zambia. In Zambia only 9-12% in the variation of official catchment populations can be explained by the estimates whereas in England 23-26% are explained. In both countries, travel-time based catchment population estimates slightly outperform straight-line distance based ones. Although the derived catchment population estimates only explain a limited share of the variation in facility headcounts, they

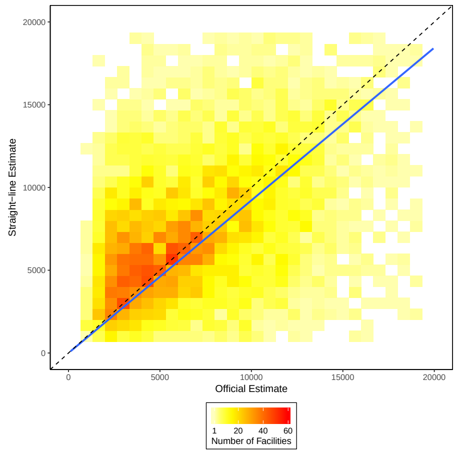
[12]Note that GPs in England may have several branches. In such cases, the catchment population for each branch is determined first. Then a GP's catchment population is computed as the sum of catchment populations over all branches.

are helpful because they allow for the inclusion of facilities without headcounts into subsequent analyses.

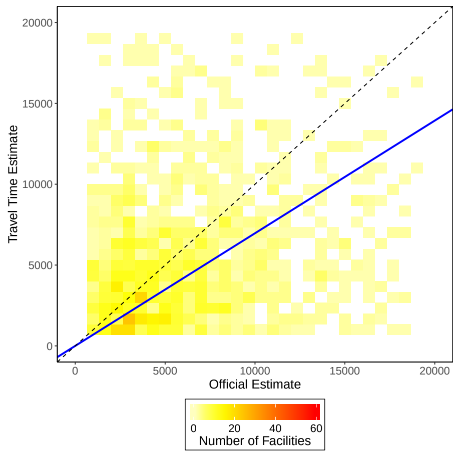
In principle, there are two reasons why estimates may differ from headcounts. First, official facility catchment areas may not correspond to estimated catchment areas and headcounts are limited to official catchment areas. Without a map of the official catchment areas, it is unfortunately not possible to assess this. Second, even if catchment areas were identical, estimates of their population would likely differ between the different considered sources due to their fundamentally different approaches to population estimation. The lower explanatory power of both methods in Zambia is likely due to lower data quality. Gridded Population of the World data is based on census data for larger administrative units in Zambia (constituencies - third-tier administrative division) than in England (output areas - sixth-tier administrative division). Additionally, registered patient numbers in England are likely to be more reliable than facility headcounts in Zambia, and travel time data is likely to be more accurate for England as well (due to higher accuracy and completeness of both Google and OpenStreetMap). Finally, higher internal migration and higher population growth in Zambia may also be reasons for the lower explanatory power. This is because the census data underlying the Gridded Population of the World data for Zambia and England is from 2010 and 2011, respectively, while official catchment population estimates are from 2017 and 2014, respectively.



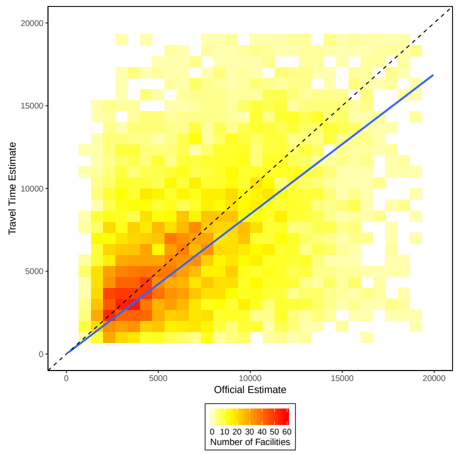
(a) Zambia - Straight-line distance



(b) England - Straight-line distance



(c) Zambia - Travel time



(d) England - Travel time

Figure 6.3: Comparison of catchment population estimates across methods and countries

Official estimates in Zambia are defined as facility headcounts as reported by facility staff in the 2017 EQUIP health facility census. In England, official estimates refer to the numbers of registered patients. Samples are limited to facilities where official estimates are available. Dashed lines are the 45-degree lines. Blue lines indicate linear regression lines. 156 and 186 facilities are outside the range of the plots and in Zambia and England, respectively.

6.3 Results

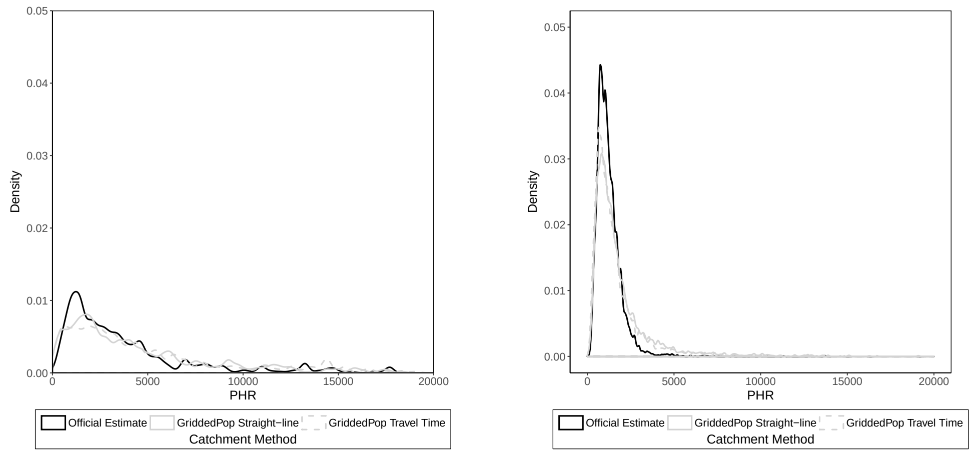
Figure 6.4(a) plots the distribution of PHRs across primary care facilities in Zambia. Facilities are weighted by their catchment population and the sample is restricted to facilities with GPS coordinates, catchment population headcounts, and staffing information. It comprises 1420 out of 2490 primary care facilities in Zambia providing care for 12 million people out of a total population of 16.59 million[13]. A distribution plot that also contains facilities without catchment population headcounts can be found in the appendix (see figure E.1). Results are very similar. The black line shows the distribution based on facility headcounts of catchment populations and the grey lines indicate the corresponding distributions based on estimated catchment populations. Independent from the source of catchment population estimates, a large variation of relative staffing across the population is observed. The long right tail of the distribution stands out. Based on official estimates, 10% of the population live in areas where the PHR is below 850. At the other extreme, 10% of the population live in areas where this ratio exceeds 8133. On average, the PHR is 3695.

The picture is very different in England. Figure 6.4(b) shows that there is hardly any right tail and PHRs vary much less across the population. Not surprisingly, the average PHR is also much lower. The mean ratio is 1238, and the 10th and 90th percentile are 564 and 2012, respectively.

So far, health workers have been defined as all medical staff at the included facilities[14]. Staffing imbalances in Zambia are even more extreme when restricting the focus to high-skilled medical staff. As figure 6.5(a) shows, the distribution of the ratio of population to doctors, nurses, and midwives has a slightly thicker tail. The ratio of the 90th and the 50th percentile of the distribution amounts to 3.3 while it was only 3.0 under

[13]Source: World Bank 2016.

[14]In line with the WHO's Global Observatory data medical staff includes nursing and midwifery personnel, dentistry personnel, pharmaceutical personnel, laboratory health workers, medical assistants, community and traditional health workers, biomedical engineers, surgical workforce and skilled health personnel.



(a) Zambia

(b) England

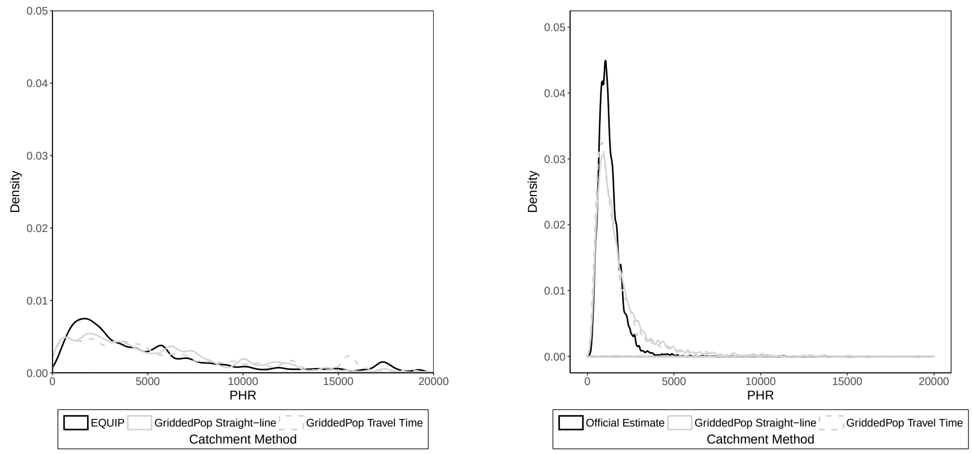
Figure 6.4: Distribution of PHRs across primary care facilities
Official estimates in Zambia are defined as facility headcounts as reported by facility staff in the 2017 EQUIP health facility census. In England, official estimates refer to the numbers of registered patients. Samples are restricted to facilities where official estimates are available. Health workers include all medical staff. Facilities are weighted by catchment population.

the more general definition of health workers[15]. Figure 6.5(b) reveals that the distribution for England hardly differs from the one in the previous figure 6.4(b). Indeed, the ratio of the 90th to the 50th percentile is almost identical[16]. This is largely because unlike in Zambia, in England most medical staff is high-skilled.

Interestingly, mapping the access to health workers across space reveals a lot of local variation in staffing across facilities in Zambia. Figure 6.6(a) shows a heat map of PHRs in Zambia. The area around each facility is colored according to the facility PHR. Shades of green indicate low PHRs and shades of red high PHRs. District borders are drawn as black lines. The map is very spotty and within most districts a relatively wide

[15]Since figure 6.5(a) does not include facilities without any high-skilled medical staff (because their PHR is infinity), this is actually an underestimate of the increase in the thickness of the right tail.

[16]The ratio of the 90th to the 50th percentile in Engalnd amounts to 1.9 considering all medical staff and 1.8 considering only high-skilled staff.)



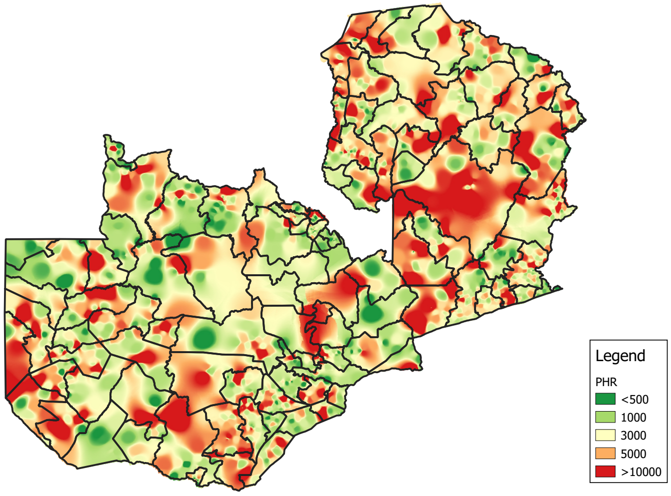
(a) Zambia (b) England

Figure 6.5: Distribution of population per high-skilled health worker across primary care facilities
Official catchment population estimates in Zambia are defined as facility headcounts as reported by facility staff in the 2017 EQUIP health facility census. In England, official catchment population estimates refer to the numbers of registered patients. Samples are restricted to facilities where official estimates are available. High-skilled health workers are defined as doctors, nurses, and midwives. Facilities are weighted by catchment population.

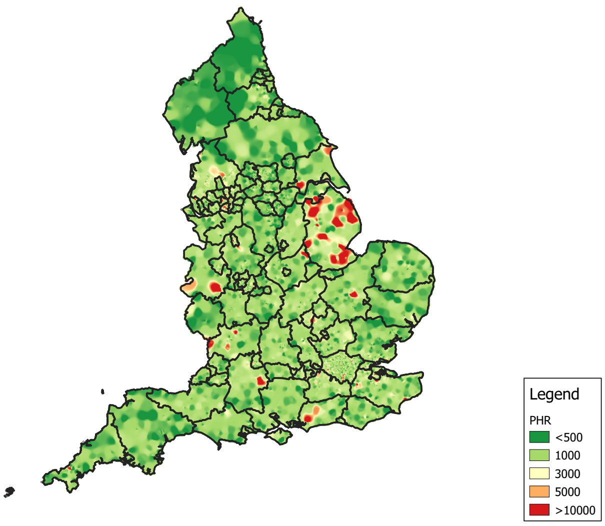
color spectrum is observed. A spatial decomposition of the cross-facility variance in PHRs in Zambia confirms this impression. The within-district standard deviation is 4595 whereas the cross-district standard deviation is only 2046. In England, on the contrary, staffing is much more balanced across space as already shown. Figure 6.6(b) confirms this once again.

Similarly in line with the findings from chapter 3, variation in facility remoteness can only explain a small share of the overall variation in staffing. Figure 6.7 shows that within each quartile of facility remoteness, as measured by population density (GPW v4) within a circle of 3km radius around the facility, there is a lot variation in PHRs[17].

[17] Results are vary similar when using alternative measures of facility remoteness. See figures E.2 and E.3



(a) Zambia



(b) England

Figure 6.6: Heat maps of PHRs
Samples are restricted to facilities where official catchment population estimates are available. Health workers include all medical staff. Black lines indicate district and county borders, respectively.

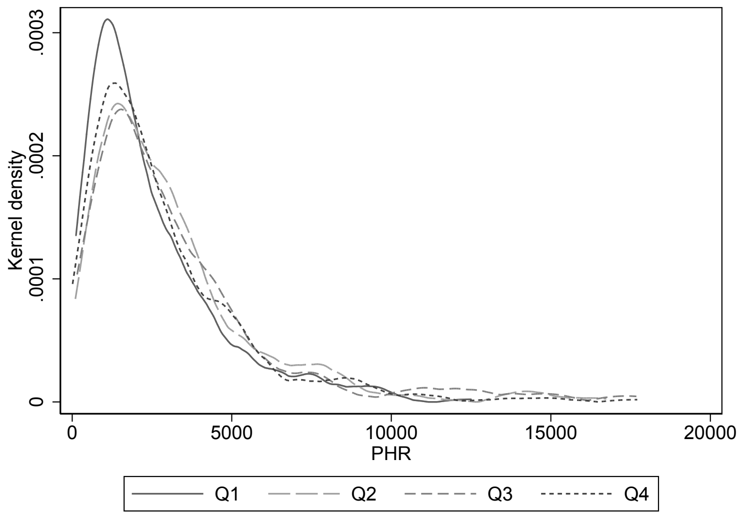


Figure 6.7: PHR distribution by quartile of population density in Zambia
Population density is measured as the density within a circle of 3km radius around a health facility based on data from the Gridded Population of the World (v4).

6.4 Discussion

The distributional patterns of health workers across public primary care facilities in Zambia and England are similar to those of teachers across public primary schools in developing and developed countries. While variation in PHRs across facilities is large in Zambia, there is little variation in England. Additionally, only a small share of the PHR variation in Zambia can be explained by facility remoteness. This suggests that not only misallocation of teachers, but also misallocation of health workers could be a serious issue in developing countries.

It remains a question for future research whether the results from Zambia can be generalized to other developing countries and to what extent the observed PHR variation is indeed indicative of health worker misallocation. Given the overall lack of health resources in developing countries and the large share of the health care budget that is typically spent on health worker salaries in these countries, the importance of this question cannot be overstated.

Chapter 7

Conclusion

This thesis constitutes a comparative analysis of the distribution of teachers across public primary schools in 86 countries. I build a new global data set comprising the universe of public primary schools in 70 countries, and subsamples in 16 countries. In line with existing evidence, I show that average PTRs in public primary schools are negatively correlated with per capita income across countries. Additionally, I present three new stylized facts consistent with teacher misallocation across schools. First, I document that the within-country variation in PTRs is higher in lower income countries. This negative correlation between PTR variation and per capita income is also found within countries over time. Second, I show in cross-country regressions and cross-district regressions that aggregate educational attainment and PTR variation are negatively correlated â even after controlling for differences in per capita income, population, and aggregate PTR. Third, I find that only a small share of PTR variation within developing countries can be explained by differences in school remoteness, as measured by population density, nighttime luminosity, or travel time to the closest city.

In order to assess to what extent teachers are misallocated in developing countries, I develop a theoretical model and calibrate it. In simulations of three different counterfactual teacher distributions I find that aggregate grade promotion gains from teacher reallocation would be substantial in many developing countries. For example, implementing a simple rule-based teacher allocation system that restricts PTRs by an achievable upper bound would increase promotion rates by 1 percentage point in Zambia and up to 4 percentage points in India. This corresponds to an additional year of education for 1% of primary school-aged children in Zambia and 4% in India. With 61% of children between the ages 5 and 14 worldwide living in low- and lower-middle-income countries[1], the implications of these findings are far-reaching.

They also raise the question why teachers are suboptimally distributed. In a case study in collaboration with the Zambian Ministry of General Education I show that lack of managerial capacity and weak enforcement are key factors. Evidence from other countries points in the same direction. Hence, the presented findings call for investment in managerial capacity and enforcement mechanisms in the public sector of developing countries. This is particularly the case because other sectors than education are likely to be similarly plagued by human resource misallocation. A comparison of the health worker distribution across public primary care facilities in Zambia and England reveals similar patterns as those observed for teachers across public primary schools, thus suggesting that the findings from this thesis apply beyond the education sector.

[1]Figure based on 2015 data from the UNESCO Institute for Statistics and World Bank International Comparison Program Database

References

Agarwal, Siddhant; Kayina, Athisii; Mukhopadhyay, Abhiroop; Reddy, Anugula. 2016. Redistributing Teachers using Local Transfers. ISI Discussion Paper 16-08. Delhi: Indian Statistical Institute.

Akhtari M.; Moreira D.; Trucco L. 2017. Political Turnover, Bureaucratic Turnover, and the Quality of Public Services. Working Paper.

Angrist, Joshua D.; Lavy, Victor. 1999 Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement. The Quarterly Journal of Economics 114 (2), 533-575.

Angrist, Joshua D. ; Lavy, Victor; Leder-Luis, Jetson; Shany, Adi. 2017. Maimonides Rule Redux. NBER Working Paper No. 23486.

Appiah-Denkyira, Ebenezer; Herbst, Christopher H.; Soucat, Agnes; Lemiere, Christophe; Saleh, Karima. 2013. Towards Interventions in Human Resources for Health in Ghana : Evidence for Health Workforce Planning and Results. Directions in Development–Human Development. Washington, DC: World Bank.

Ashraf, Nava; Bandiera, Oriana; Lee, Scott. 2018. Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the Delivery of Public Services. Working Paper.

Asim, Salman; Chimombo, Joseph P. G.; Chugunov, Dmitry; Gera, Ravinder Madron Casley. 2017. Moving teachers to Malawi's remote communities: a data-driven approach to teacher deployment (English). Policy Research Working Paper Series 8253. Washington, DC: World Bank Group.

Baker, Carl. 2016. Population estimates & GP registers: why the difference. House of Commons Library. Retrieved from *https://commonslibrary.parliament.uk/social-policy/health/population-estimates-gp-registers-why-the-difference/* on 13/07/2018.

Bandiera, O.; Larcinese, V.; Rasul, I. 2010. Heterogeneous Class Size Effects: New Evidence from a Panel of University Students. The Economic Journal 120, 1365-1398.

Banerjee, A.; Deaton, A.; Duflo, E. 2004. Health care delivery in rural Rajasthan. Poverty Action Lab working paper No. 4.

Barrios, Andres; Bovini, Giulia. 2017. It's Time to Learn: Understanding the Differences in Returns to Instruction Time. Working paper.

Beteille, T. 2009. Absenteeism, Transfers and Patronage: The Political Economy of Teacher Labour Markets in India. PhD thesis, Stanford University.

Bold, Tessa; Filmer, Deon P.; Martin, Gayle; Molina, Ezequiel; Rockmore, Christophe; Stacy, Brian William; Svensson, Jakob; Wane, Waly. 2017. What do teachers know and do? Does it matter? Evidence from primary schools in Africa. Policy Research Working Paper Series 7956. Washington, DC: World Bank Group.

Busso, M.; Madrigal, L.; Pages, C. 2013. Productivity and resource misallocation in Latin America. The B.E. Journal of Macroeconomics

13(1), 903-932.

Chaudhury, N.; Hammer, J.; Kremer, M.; Muralidharan, K.; Halsey Rogers, F. 2006. Missing in action: Teacher and health worker absence in developing countries. Journal of Economic Perspectives 20 (1), 91-116.

Checchi, Daniele; De Paola, Maria. 2017. The Effect of Multigrade Classes on Cognitive and Non-Cognitive Skills: Causal Evidence Exploiting Minimum Class Size Rules in Italy. IZA DP No. 11211. Bonn: Institute of Labor Economics.

Chin, Aimee. 2005. Can redistributing teachers across schools raise educational attainment? Evidence from Operation Blackboard in India. Journal of Development Economics 78 (2), 384-405.

Cummings, Clare and Tahirou, Ali Bako M. 2016. Collective action and the deployment of teachers in Niger: a political economy analysis. ODI briefing. London: Overseas Development Institute.

Dal Bo, Ernesto; Finan, Frederico; Rossi, Martin A. 2013. Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service. The Quarterly Journal of Economics 128 (3), 1169-1218.

Das, J.; Hammer, J. 2007. Money for nothing: The dire straits of medical practice in Delhi, India. Journal of Development Economics 83 (1), 1-36.

Das, Jishnu; Hammer, Jeffrey; Leonard, Kenneth. 2008. The Quality of Medical Advice in Low-Income Countries. Policy Research Working Paper No. 4501. Washington, DC: World Bank.

De Ree, Joppe; Muralidharan, Karthik; Pradhan, Menno; Rogers, Halsey. 2018. Double for Nothing? Experimental Evidence on an Unconditional Teacher Salary Increase in Indonesia. The Quarterly Journal of Economics

133 (2), 993-1039.

Diompy, Danty Patrick. 2014. De la mobilite de carrieres du personnel enseignant dans le moyen secondaire au Senegal : perceptions des acteurs. Master thesis. Universite Cheikh Anta Diop de Dakar, Faculte des Sciences et Technologies, de l'Education et de la Formation.

Duflo, Esther; Dupas, Pascaline; Kremer, Michael. 2015. School governance, teacher incentives, and pupil-teacher ratios: Experimental evidence from Kenyan primary schools. Journal of Public Economics (123), 92-110.

Duflo, Esther; Hanna, Rema; Ryan, Stephen P. 2012. Incentives Work: Getting Teachers to Come to School. The American Economic Review 102 (4), 1241-1278.

Dussault, G., and Franceschini, M. C. 2006. Not Enough There, Too Many Here: Understanding Geographical Imbalances in the Distribution of the Health Workforce. Human Resources for Health 4 (12).

Fagernas, Sonja and Pelkonen, Panu. 2012. Preferences and skills of Indian public sector teachers. IZA Journal of Labor & Development, 1 (3).

Fagernas, Sonja and Pelkonen, Panu. 2017. Where's the Teacher? How Teacher Workplace Segregation Impedes Teacher Allocation in India. IZA DP No. 10595. Bonn: Institute of Labor Economics.

Ferrinho, P.; Siziya, S.; Goma, F.; Dussault, G. 2011. The Human Resource for Health Situation in Zambia: Deficit and Maldistribution. Human Resources for Health 9 (30).

Franck, R.; Rainer, I. 2012. Does the Leader's Ethnicity Matter? Ethnic Favoritism, Education, and Health in Sub-Saharan Africa. American

Political Science Review 106(2), 294-325.

Glewwe, P.; Kremer, M.; Moulin, S.; Zitzewitz, E. 2004. Retrospective vs. prospective analyses of school inputs: the case of flip charts in Kenya. Journal of Development Economics 74 (1), 251-268.

Glewwe, Paul; Muralidharan, Karthik. 2016. Improving Education Outcomes in Developing Countries: Evidence, Knowledge Gaps, and Policy Implications. Handbook of the Economics of Education 5, 653-743.

Hedges, John. 2002. The importance of posting and interaction with the education bureaucracy in becoming a teacher in Ghana. International Journal of Educational Development 22 (3-4), 353-366.

Hoxby, Caroline M. 2000. The Effects of Class Size on Student Achievement: New Evidence from Population Variation. The Quarterly Journal of Economics 115 (4), 1239-1285.

Hsieh, Chang-Tai; Klenow, Peter J. 2009. Misallocation and Manufacturing TFP in China and India. Quarterly Journal of Economics 124 (4), 1403-1448.

IIEP- Pole de Dakar. 2016. More effective teacher allocation in Africa. Polemag - IIEP Pole de Dakar Information Magazine 24, 8-13.

Jacob, Verghese; Kochar, Anjini; Reddy, Suresh. 2008. School Size and Schooling Inequalities. Stanford Center on Global Poverty and Development Working Paper 354.

Kramon, Eric; Posner, Daniel N. 2016. Ethnic Favoritism in Education in Kenya. Quarterly Journal of Political Science 11 (1), 1-58.

Krueger, Alan B. 1999. Experimental Estimates of Education Production Functions. The Quarterly Journal of Economics 114 (2), 497-532.

Lavy, Victor. 2015. Do Differences in Schools' Instruction Time Explain International Achievement Gaps? Evidence from Developed and Developing Countries. The Economic Journal 125 (588), F397-F424.

Lemiere, Christophe; Herbst, Christopher H.; Dolea, Carmen; Zurn, Pascal; Soucat, Agnes. 2013. Rural-Urban Imbalance of Health Workers in Sub-Saharan Africa. In "The Labor Market for Health Workers in Africa: A New Look at the Crisis". Washington, DC: World Bank.

Lemiere, Christophe; Herbst, Christopher H.; Jahanshahi, Negda; Smith, Ellen; Soucat, Agnes. 2011. Reducing Geographical Imbalances of Health Workers in Sub-Saharan Africa : A Labor Market Perspective on What Works, What Does Not, and Why. World Bank Working Paper No. 209. Africa Human Development Series. World Bank.

Lemos, Renata; Scur, Daniela. 2016. Developing Management: An expanded evaluation tool for developing countries. RISE Working Paper 7.

Miller, Grant; Babiarz, Kim. 2014. Pay-for-Performance Incentives in Low- and Middle-Income Country Health Programs. In "Encyclopedia of Health Economics". Elsevier.

Mingat, Alain; Tan, Jee-Peng; Sosale, Shobhana. 2003. Tools for Education: Policy Analysis. Washington, DC: World Bank.

Ministere de l'economie, de la planification, et de l'amenagement du territoire, Republique du Cameroun. 2013. Document de Strategie du Secteur de l'Education et de la Formation (2013-2020).

Ministry of Education, Science, Vocational Training and Early Education, Government of the Republic of Zambia. 2011. Education Sector National Implementation Framework III 2011-2015.

Ministry of Primary and Mass Education, Government of the People's Republic of Bangladesh. 2015. Third Primary Education Development Program (PEDP-3) - Revised.

Mulkeen, Aidan. 2010. Teachers in Anglophone Africa : Issues in Teacher Supply, Training, and Management. Development Practice in Education. Washington, DC: World Bank.

Munga, M., and Maestad, O. 2009. Measuring Inequalities in the Distribution of Health Workers: The Case of Tanzania. Human Resources for Health 7 (4).

Muralidharan, Karthik; Das, Jishnu; Holla, Alaka; Mohpal, Aakash. 2017. The fiscal cost of weak governance: Evidence from teacher absence in India. Journal of Public Economics 145, 116-135.

Muralidharan, K.; Sundararaman, V. 2011. Teacher performance pay: experimental evidence from India. Journal of Political Economy 119 (1), 39-77.

Muralidharan, K.; Sundararaman, V. 2013. Contract Teachers: Experimental Evidence from India. NBER Working Paper No. 19440.

Panjabi, Raj. 2016. Four difficult truths highlighted by the Ebola epidemic. TED. Retrieved from *https://ideas.ted.com/four-difficult-truths-highlighted-by-the-ebola-epidemic/* on 24/06/2018.

Pugatch, Todd; Schroeder, Elizabeth. 2014. Incentives for teacher relocation: Evidence from the Gambian hardship allowance. Economics of Education Review 41, 120-136.

Ramachandran, Vimala; Beteille, Tara; Linden, Toby; Dey, Sangeeta; Goyal, Sangeeta; Goel Chatterjee, Prerna. 2018. Getting the Right Teachers into the Right Schools : Managing India's Teacher Workforce. World Bank Studies. Washington, DC: World Bank.

Restuccia, Diego; Rogerson, Richard. 2017. The Causes and Costs of Misallocation. Journal of Economic Perspectives 31 (3), 151-74.

Rivkin, Steven G.; Schiman, Jeffrey C. 2015. Instruction time, classroom quality, and academic achievement. The Economic Journal 125 (588) F425-F448.

Scheffler, Richard M.; Campbell, James; Cometto, Giorgio; Maeda, Akiko; Liu, Jenny; Bruckner, Tim A.; Arnold, Daniel R.; Evans, Tim. 2018. Forecasting imbalances in the global health labor market and devising policy responses. Human Resources for Health 16 (5).

Sharma, Rashmi; Ramachandran, Vimala. 2009. The Elementary Education System in India: Exploring Institutional Structures, Processes and Dynamics. New Delhi: Routledge India.

Steiner-Khamsi, Gita. 2010. Teacher Recruitment, Development and Retention. UNICEF ESARO: 3-country study in Lesotho, Swaziland, Malawi. Funded by UNICEF ESARO, Nairobi/Kenya.

Sow, Soule. 2015. Are Cities Preferred to Villages? Estimating Location Preference in A Developing Country. Working paper.

Weiss, Daniel; Nelson, A.; Gibson, H.S.; Temperley,W.; Peedell, S.; Lieber, A.; Hancher, M.; Poyart, E.; Belchior, S.; Fullman, N.; Mappin, B.; Dalrymple, U.; Rozier, J.; Lucas, T.C.D.; Howes, R.E.; Tusting, L.S.; Kang, S.Y.; Cameron, E.; Bisanzio, D.; Battle, K.E.; Bhatt, S.; Gething, P.W. 2018. A global map of travel time to cities to assess inequalities in accessibility in 2015. Nature.

World Bank. 2008. Teacher Employment and Deployment in Indonesia: Opportunities for Equity, Efficiency and Quality Improvement. Washington, DC: World Bank.

World Bank. 2018. World Development Report 2018: Learning to Realize Education's Promise. Washington, DC: World Bank.

UNESCO. 2006. Teachers and Educational Quality: Monitoring Global Needs for 2015. Montreal: UNESCO Institute for Statistics.

Appendices

Appendix A

Appendix to Chapter 1: Introduction

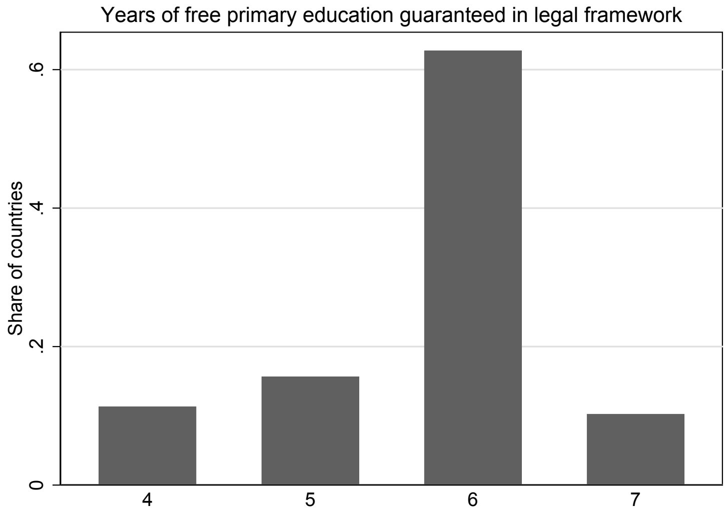


Figure A.1: Years of free primary education guaranteed in legal framework across countries
Source: UNESCO Institute for Statistics and World Bank International Comparison Program Database. Data from 2015. Sample size: 185 countries.

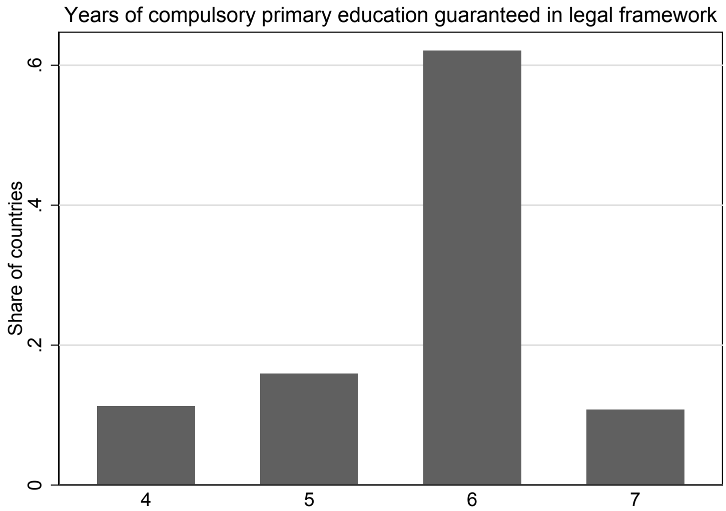


Figure A.2: Years of compulsory primary education guaranteed in legal framework across countries
Source: UNESCO Institute for Statistics and World Bank International Comparison Program Database. Data from 2015. Sample size: 195 countries.

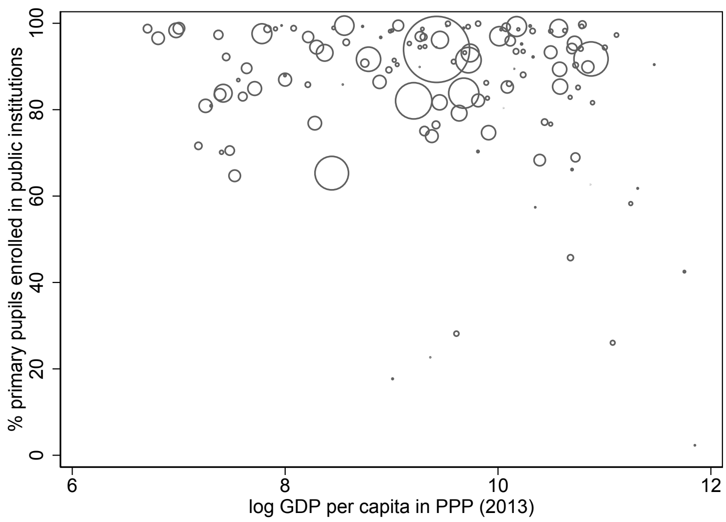


Figure A.3: Share of primary school pupils in public institutions and income across countries
Source: UNESCO Institute for Statistics and World Bank International Comparison Program Database. Data from 2013. Sample size: 128 countries. Marker size indicates size of the primary school aged population (ages 5-14).

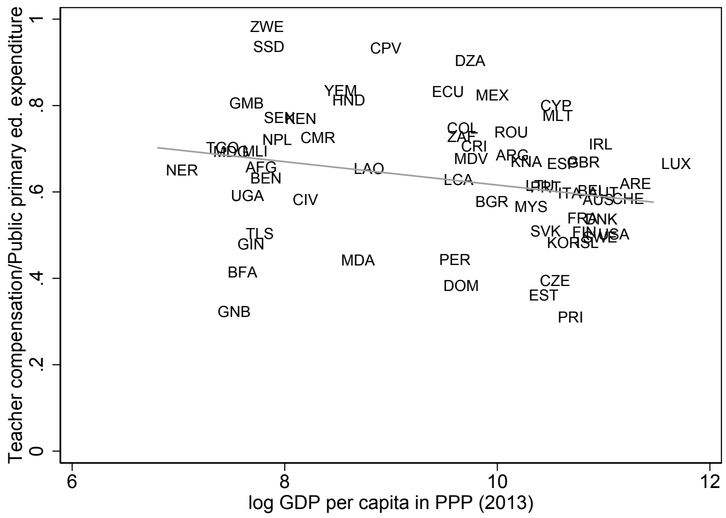


Figure A.4: Government expenditure on teacher compensation and income across countries
Source: UNESCO Institute for Statistics and World Bank International Comparison Program Database. Data from 2013. Sample size: 64 countries.

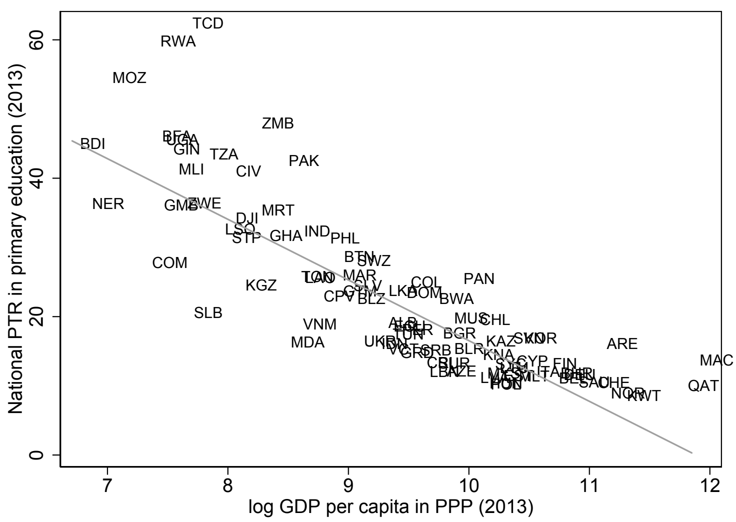


Figure A.5: National PTR in primary education and per capita income across countries
Source: UNESCO Institute for Statistics and World Bank International Comparison Program Database. Data from 2013. Sample size: 83 countries.

Appendix B

Appendix to Chapter 2: Theoretical framework for the analysis of teacher allocation

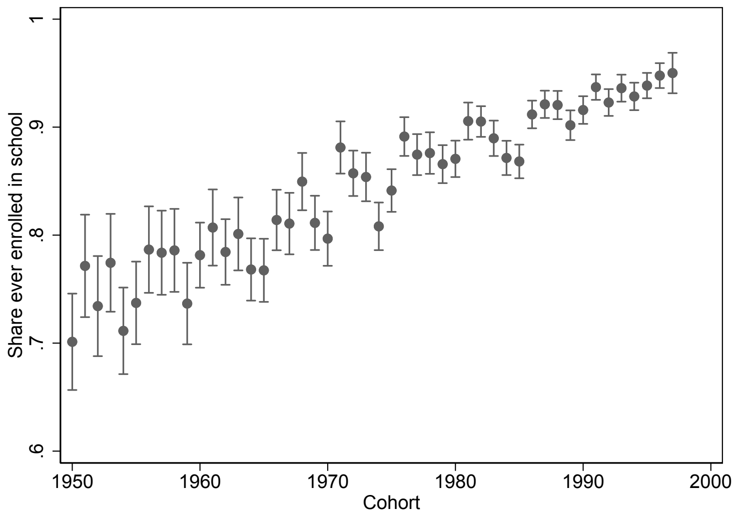
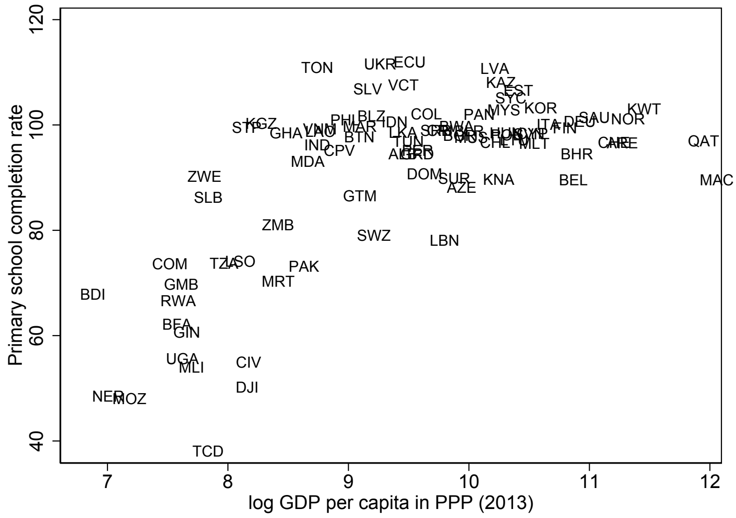
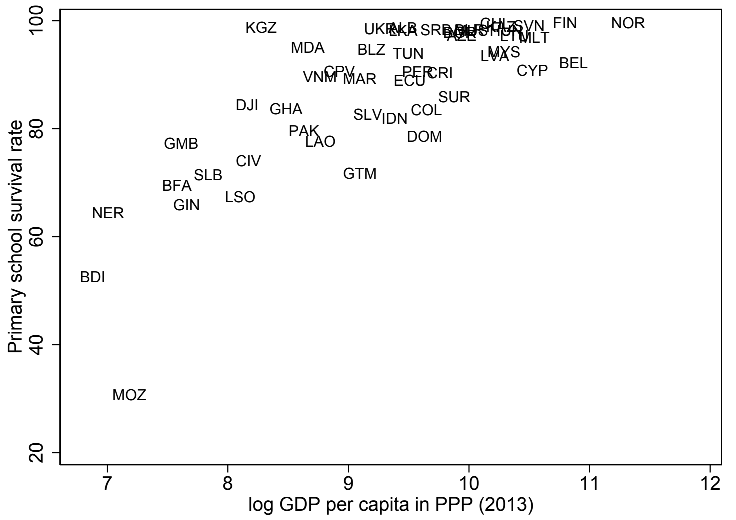


Figure B.1: Primary school enrollment in Africa

Source: Afrobarometer Round 6 (2014-2015). Sample: 36 countries.



(a) Primary school completion rate



(b) Survival to last grade of primary education

Figure B.2: Primary school completion and survival across countries by per capita income

Source: UNESCO Institute for Statistics and World Bank International Comparison Program Database (2013). Sample size: 83 and 49 countries, respectively.

Appendix C

Appendix to Chapter 3: Cross-country comparison of the distribution of teachers across public primary schools

Table C.1: Core data sources

Country	State/Province	Data type	Collection method	Data source	Date obtained	Share w/o PTR
Antigua and Barbuda		Census	Download	Educational Statistical Digest 2012; Ministry of Education, Sports, Youth and Gender Affairs; Antigua and Barbuda; Retrieved from: http://www.education.gov.ag/#	07.08.2017	0%
Argentina		Census	Download	Direccion Nacional de Informacion y Estadistica Educativa; Ministerio de Educacion y Deportes; Argentina; Retrieved from: http://portales.educacion.gov.ar/diniece/2016/08/24/bases-de-datos-por-escuela-con-id/	22.01.2017	1%
Australia[1]		Census	Data request	Australian Curriculum, Assessment and Reporting Authority	25.04.2017	$<$ 1%
Austria	Burgenland	Census	Data request	Landesschulrat fuer Burgenland	18.04.2017	0%
Austria	Niederoesterreich	Census	Data request	Landesschulrat fuer Niederoesterreich	25.04.2017	0%
Austria	Oberoesterreich	Census	Data request	Landesschulrat fuer Oberoesterreich	31.03.2017	0%
Austria	Steiermark	Census	Data request	Landesschulrat fuer Steiermark	02.05.2017	0%
Belgium	Flanders	Census	Data request	Education and Training; Flemish Community of Belgium	13.12.2017	0%
Benin		Survey	Data request	Programme d'analyse des systemes educatifs de la confemen (PASEC) 2014	07.07.2017	N/A
Bhutan		Census	Download	Annual Education Statistics 2015; Ministry of Education; Royal Government of Bhutan; Retrieved from: http://www.education.gov.bt/statistic	22.01.2017	$<$ 1%
Botswana		Census	Download	Ministry of Education and Skills Development; Botswana; Retrieved from: http://www.gov.bw/en/Ministries–Authorities/Ministries/Ministry-of-Education-and-Skills-Development/Schools/Public-Primary-Schools/	18.11.2016	$<$ 1%
Brazil		Census	Download	Instituto Nacional de Estudos e Pesquisas Educacionais Anisio Teixeira; Ministerio da Educacao; Brazil; Retrieved from: http://dados.gov.br/dataset/microdados-do-censo-escolar	04.11.2016	0%
Burkina Faso		Census	Data request	Ministere de l'Education Nationale et de l'Alphabetisation, Burkina Faso	15.03.2018	0%
Burundi		Survey	Data request	Programme d'analyse des systemes educatifs de la confemen (PASEC) 2014	07.07.2017	N/A
Cambodia		Census	Download	Ministry of Education, Youth and Sports; Cambodia; Retrieved from: https://opendevelopmentcambodia.net/dataset/?id=school-of-cambodia-2012	28.03.2018	9%
Cameroon		Survey	Data request	Programme d'analyse des systemes educatifs de la confemen (PASEC) 2014	07.07.2017	N/A
Canada	New Brunswick	Census	Data request	Department of Education and Early Childhood Development; New Brunswick	15.06.2017	0%
Canada[2]	Ontario	Census	Data request	Ministry of Education; Ontario	20.06.2017	18%
Cape Verde		Census	Download	Anuario da Educacao 2014/2015; Ministerio da Educacao e Desporto; Cape Verde; Retrieved from: http://www.minedu.gov.cv/index.php?option=com_jdownloads&view=summary&id=913:anuario-da-educacao-ano-letivo-2014-2015&catid=4&Itemid=574	23.11.2016	5%
Chad		Survey	Data request	Programme d'analyse des systemes educatifs de la confemen (PASEC) 2014	07.07.2017	N/A
Chile		Census	Download	Centro de Estudios; Ministerio de Educacion; Gobierno de Chile; Retrieved from: http://centroestudios.mineduc.cl/tp_modulos/tpm_seccion/contVentana.php?cc=2179	07.11.2016	0%
Colombia		Census	Download	Investigacion de Educacion Formal; Departamento Administrativo Nacional de Estadistica; Colombia; Retrieved from: http://microdatos.dane.gov.co/index.php/catalog/MICRODATOS/about$_{c}ollection$/25/2	14.08.2018	1%
Congo, Rep.		Survey	Data request	Programme d'analyse des systemes educatifs de la confemen (PASEC) 2014	07.07.2017	N/A
Costa Rica		Census	Data request	Departamento de AnÃ¡lisis EstadÃstico; Ministerio de Educacion Publica; Costa Rica	17.05.2017	0%

Table C.1: Core data sources

Country	State/Province	Data type	Collection method	Data source	Date obtained	Share w/o PTR
Cote d'Ivoire		Survey	Data request	Programme d'analyse des systemes educatifs de la confemen (PASEC) 2014	07.07.2017	N/A
Czech Republic		Census	Data request	Ministry of Education, Youth and Sports of the Czech Republic	07.04.2017	$< 1\%$
Denmark		Census	Data request	Undervisningsministeriet; Styrelsen for It og LÃring; Center for Data og Analyse; Denmark	18.09.2017	2%
Djibouti		Census	Download	Annuaire Statistique 2014-2015; Ministere de lâEducation Nationale et de la Formation Professionnelle; Republique de Djibouti; Retrieved from: http://www.education.gov.dj/	09.01.2017	0%
Dominican Republic		Census	Data request	Instituto Dominicano de Evaluacion e Investigacion de la Calidad Educativa	13.04.2018	0%
Ecuador		Census	Data request	Ministerio de Educacion; Ecuador	03.03.2017	$< 1\%$
El Salvador		Census	Download	Ministerio de Educacion; Republica de El Salvador; Retrieved from: https://www.mined.gob.sv/index.php/estadisticas-educativas/item/6116-bases-de-centros	11.12.2016	0%
Estonia		Census	Data request	Analysis Department; Estonian Ministry of Education and Research	24.04.2017	0%
Fiji		Census	Download	Ministry of Education; Fiji; Retrieved from: http://www.education.gov.fj/	07.04.2017	1%
France		Census	Download	Ministere de l'Education nationale, de l'Enseignement Superieur et de la Recherche; France; Retrieved from: https://data.education.gouv.fr/	23.01.2017	$< 1\%$
Georgia		Census	Download	eCatalog; Education Management Information System; Georgia; Retrieved from: http://catalog.edu.ge/index.php?module=school_info&page=region_list	29.05.2017	0%
Germany	Hamburg	Census	Data request	Behoerde fuer Schule und Berufsbildung; Freie und Hansestadt Hamburg	15.05.2017	0%
Germany	Hessen	Census	Data request	Hessisches Kultusministerium	28.03.2017	0%
Germany	Schleswig-Holstein	Census	Data request	Ministerium fuer Schule und Berufsbildung des Landes Schleswig-Holstein	03.04.2017	0%
Germany	Thueringen	Census	Data request	Thueringer Ministerium fuer Bildung, Jugend, und Sport	20.03.2017	0%
Guatemala		Census	Download	Direccion de Planificacion; Ministerio de Educacion; Guatemala; Retrieved from: http://estadistica.mineduc.gob.gt/BDD/	12.11.2016	4%
Guinea-Bissau		Census	Data request	Ministerio da Educacao e do Ensino Superior; Guinea-Bissau	11.12.2017	6%
Honduras		Census	Download	Unidad de Planeamiento y Evaluacion la Gestion; Secretaria de Educacion de Honduras; Retrieved from: http://estadisticas.se.gob.hn/see/archivos_descargables.php	18.12.2017	0%
Hungary[3]		Census	Data request	Hungarian Central Statistical Office	06.03.2017	$< 1\%$
India		Census	Data request	District Information System for Education; National University of Educational Planning and Administration; India	24.11.2016	1%
Ireland		Census	Download	Department of Education and Skills; Ireland; Retrieved from: http://www.education.ie/en/Publications/Statistics/Data-on-Individual-Schools/Data-on-Individual-Schools.html	17.11.2016	0%
Jamaica		Census	Data request	Ministry of Education, Jamaica	24.02.2017	0%
Kenya		Survey	Download	UWEZO; Retrieved from: http://www.uwezo.net/publications/datasets/	21.03.2017	N/A
Kiribati		Census	Download	Digest of Education Statistics 2011; Ministry of Education; Republic of Kiribati; Retrieved from: http://prism.spc.int/reports/education	25.11.2016	0%
Kyrgyzstan		Census	Download	Ministry of Education and Science of the Kyrgyz Republic; Retrieved from: http://edu.gov.kg/ru/docs/statistics/	07.02.2017	$< 1\%$
Laos		Census	Data request	Ministry of Education and Sports; Lao People's Democratic Republic	10.04.2018	0%
Latvia		Census	Data request	Ministry of Education and Science of the Republic of Latvia	21.03.2017	0%
Liberia		Census	Download	Ministry of Education; Republic of Liberia; Retrieved from: http://moe.gov.lr/documents/	06.12.2017	5%

Table C.1: Core data sources

Country	State/Province	Data type	Collection method	Data source	Date obtained	Share w/o PTR
Libya		Census	Download	Libya National Schools Assessment 2012; REACH Initiative; Retrieved from: https://data.humdata.org/dataset/reach-libya-national-schools-assessment-2012	13.02.2017	11%
Lithuania		Census	Data request	Ministry of Education and Science of the Republic of Lithuania	06.03.2017	0%
Madagascar		Census	Data request	Madagascar Ministere de l'Education Nationale	11.03.2017	< 1%
Malawi		Census	Data request	Ministry of Education; Malawi	21.03.2017	0%
Marshall Islands		Census	Download	Education Digest 2013-2014; Ministry of Education; Republic of the Marshall Islands; Retrieved from: http://prism.spc.int/reports/education	25.11.2016	0%
Mexico		Census	Download	Sistema Nacional de Informacion de Escuelas; Secretaria de Educacion Publica; Mexico; Retrieved from: http://www.snie.sep.gob.mx/SNIESC/	02.10.2017	2%
Moldova		Census	Data request	National Bureau of Statistics of the Republic of Moldova	10.05.2017	0%
Mongolia		Census	Data request	National Statistical Office of Mongolia	12.01.2017	4%
Mozambique		Census	Data request	Ministerio da Educacao e Desenvolvimento Humano; Republica da Mocambique	08.05.2017	0%
Netherlands		Census	Download	Dienst Uitvoering Onderwijs; Ministerie van OCW; Retrieved from: https://www.duo.nl/open_onderwijsdata/databestanden/po/	06.02.2017	< 1%
New Zealand		Census	Download	Education Counts; Ministry of Education; New Zealand Government; Retrieved from: http://www.educationcounts.govt.nz/statistics/schooling/	02.11.2016	< 1%
Niger		Survey	Data request	Programme d'analyse des systemes educatifs de la confemen (PASEC) 2014	07.07.2017	N/A
Norway		Census	Data request	Norwegian Directorate for Education and Training; Department of Statistics	08.02.2017	< 1%
Pakistan	Balochistan	Census	Download	Balochistan EMIS; Retrieved from: http://emis.gob.pk/views/Reports/Reports/SchoolSearchPublic.aspx	02.09.2018	12%
Pakistan	Punjab	Census	Download	Department of School Education; Government of Punjab; Retrieved from: http://www.pesrp.edu.pk/datacenter#district_ranking	19.05.2017	3%
Pakistan	Sindh	Census	Download	Education and Literacy Department; Government of Sindh; Retrieved from: http://www.rsu-sindh.gov.pk/downloads/schoolSearch.php	23.05.2017	< 1%
Palau		Census	Download	2011 Statistical Yearbook; Ministry of Education; Republic of Palau; Retrieved from: http://prism.spc.int/reports/education	25.05.2018	0%
Papua New Guinea		Census	Download	Department of Education; Papua New Guinea; Retrieved from: http://www.education.gov.pg/quicklinks/wms/school-profile.html	06.12.2016	< 1%
Paraguay		Census	Download	Ministerio de Educacion y Ciencias; Paraguay; Retrieved from: http://datos.mec.gov.py/data	24.01.2017	16%
Peru		Census	Download	Censo Escolar 2016; Ministerio de Educacion; Peru; Retrieved from: http://escale.minedu.gob.pe/uee/-/document_library_display/GMv7/view/2979785	07.11.2016	0%
Philippines		Census	Download	Department of Education; Republic of the Philippines; Retrieved from: http://www.deped.gov.ph/datasets	22.11.2016	2%
Poland		Census	Download	Centrum Informatyczne Edukacji; Poland; Retrieved from: https://cie.men.gov.pl/sio-strona-glowna/podstawowe-informacje-dotyczce-wykazu-szko-i-placowek-owiatowych/wykaz-wg-typow/	29.01.2018	26%
Puerto Rico		Census	Download	U.S. Department of Education; National Center for Education Statistics; Common Core of Data (CCD); Public Elementary/Secondary School Universe Survey CCD School Data 2014-15; Retrieved from: http://nces.ed.gov/ccd/	10.07.2017	< 1%
Saint Kitts and Nevis		Census	Download	Statistical Digest 2013-2014; St. Kitts and Nevis Ministry of Education; Retrieved from: http://www.moeskn.org/	08.12.2016	0%

Table C.1: Core data sources

Country	State/Province	Data type	Collection method	Data source	Date obtained	Share w/o PTR
Saint Lucia		Census	Download	Education Statistical Digest 2015; Ministry of Education, Human Resource Development and Labour; Government of St. Lucia; Retrieved from: http://education.govt.lc/publications/	08.12.2016	0%
Saint Vincent and the Grenadines		Census	Download	Education Statistical Digest of St. Vincent & the Grenadines 2014-2015; Retrieved from: http://www.education.gov.vc/education/	25.05.2018	0%
Samoa		Census	Download	Education Statistical Digest 2015; Ministry of Education, Sports, and Culture; Samoa; Retrieved from: http://prism.spc.int/reports/education	25.11.2016	0%
Senegal		Survey	Data request	Programme d'analyse des systemes educatifs de la confemen (PASEC) 2014	07.07.2017	N/A
Seychelles		Census	Download	Education Statistics 2012; Ministry of Education; Republic of Seychelles; Retrieved from: http://www.education.gov.sc/Pages/statistics.aspx	24.11.2016	0%
South Africa		Census	Download	Education Management Information System; National Department of Basic Education; South Africa; Retrieved from: http://www.education.gov.za/Programmes/EMIS.aspx	02.11.2016	4%
South Sudan		Census	Download	Education Management Information System; South Sudan; Retrieved from: http://www.southsudanemis.org/data	04.11.2016	$< 1\%$
Sudan	Karthoum	Census	Download	Ministry of Education; Sudan; Retrieved from: http://moekh.gov.sd/	10.02.2017	23%
Suriname		Census	Data request	Ministerie van Onderwijs, Wetenschap un Cultuur; Suriname	23.08.2017	$< 1\%$
Swaziland		Census	Data request	Ministry of Education and Training; Swaziland	30.11.2016	0%
Sweden		Census	Download	SiRiS; National Agency for Education; Sweden; Retrieved from: http://siris.skolverket.se/siris/	25.05.2018	$< 1\%$
Tanzania		Census	Download	President's Office; Regional Administration and Local Government; The United Republic of Tanzania; Retrieved from: http://opendata.go.tz/dataset/uwiano-wa-mwalimu-kwa-wanafunzi-kwa-shule-za-msingi-za-serikali-2016	04.11.2016	9%
Togo		Survey	Data request	Programme d'analyse des systemes educatifs de la confemen (PASEC) 2014	07.07.2017	N/A
Uganda		Census	Data request	Ministry of Education and Sports; The Republic of Uganda	23.08.2017	$< 1\%$
UK	England	Census	Download	Department for Education; England; Retrieved from: https://data.gov.uk/dataset/	07.02.2017	$< 1\%$
UK	Northern Ireland	Census	Download	Department for Education; Northern Ireland; Retrieved from: https://www.education-ni.gov.uk/articles/	07.02.2017	0%
UK	Scotland	Census	Download	Scottish Government; Retrieved from: http://www.gov.scot/Topics/Statistics/Browse/School-Education/	07.02.2017	$< 1\%$
UK	Wales	Census	Download	StatsWales; Retrieved from: https://statswales.gov.wales/Catalogue/Education-and-Skills/Schools-and-Teachers/Schools-Census/Pupil-Level-Annual-School-Census/	07.02.2017	0%
Ukraine		Census	Download	School Map of Ukraine; Retrieved from: http://cedos.org.ua/edustat/databox	29.01.2017	1%
Uruguay		Census	Download	Administracion Nacional de Educacion Publica; Uruguay; Retrieved from: http://www.anep.edu.uy/portalmonitor/servlet/buscarescuela	12.03.2018	$< 1\%$
US		Census	Download	U.S. Department of Education; National Center for Education Statistics; Common Core of Data (CCD); Public Elementary/Secondary School Universe Survey CCD School Data 2014-15; Retrieved from: http://nces.ed.gov/ccd/	10.07.2017	1%
Zambia		Census	Data request	Ministry of Education; Zambia	14.06.2016	12%

Table C.1: Core data sources

Country	State/Province	Data type	Collection method	Data source	Date obtained	Share w/o PTR

This table lists the data sources for the core data for all countries. In the last column, it also indicates the share of public primary schools for which the PTR could not be computed due to missing information. This is computed as the number of public primary school without PTR information over the total number of public primary schools listed. It is possible that for a given country the obtained list of public primary schools itself is incomplete. In this case the indicated share of schools for which the PTR could not be computed is an underestimate of the true share without PTR information.

[1] Disclaimer: The data used in this publication are sourced from the Australian Curriculum, Assessment and Reporting Authority (ACARA) and are available from ACARA in accordance with its Data Access Protocols.

[2] PTR information for schools with less than 10 teacher FTEs was not available. This accounts entirely for the share of schools without PTR information.

[3] Disclaimer: Results for Hungary have been created with the use of WTorsten_cimlista_altisk_tan_ped_2015-16.xlsx Datafile prepared upon individual request by the Hungarian Central Statistical Office (www.ksh.hu). The calculations and the conclusion are the sole intellectual products of the author Torsten Figueiredo Walter.

Table C.2: PTR time series data

Country	Time Span
Argentina	2011-2015
Bhutan	2005-2016
Cape Verde	2004, 2006-2015
Chile	2004-2016
El Salvador	2007-2010, 2013
India	2005-2006, 2008-2015
Ireland	2006-2015
Mozambique	2004-2016
Netherlands	2011-2016
New Zealand	2008-2016
Peru	2004-2016
Puerto Rico	1990-2015
Saint Vincent and the Grenadines	2011-2016
South Africa	2007-2013, 2015-2016
South Sudan	2008-2013, 2015
Sweden	1999-2016
Uganda	2013-2016
Uruguay	2011-2015
US	1990-2015
Zambia	2002, 2006-2016

This table lists the countries for which school-level pupil-teacher ratio data from the universe of public primary schools was collected for four or more years. For each country, the table indicates the years for which data was obtained. The source of the additional years of data is identical to the source provided in table C.1.

Table C.3: Regions and subregions

Country	Region definition	Regions	Subregion definition	Subregions
Antigua and Barbuda	Education Zone	4	N/A	N/A
Argentina	Province	24	N/A	N/A
Australia	State	8	N/A	N/A
Austria	State (NUTS-2)	4	Groups of Municipalities (NUTS-3)	25
Belgium	Province (NUTS-2)	6	Arrondissements (NUTS-3)	23
Bhutan	Dzongkhag	20	N/A	N/A
Botswana	District	14	N/A	N/A
Brazil	State	27	Municipality	5556
Burkina Faso	Region	13	Province	45
Cambodia	Province	24	District	186
Canada	Province	2	County/District	66
Cape Verde	Concelho	22	N/A	N/A
Chile	Region	13	Province	53
Colombia	Department	33	Municipality	1118
Costa Rica	Province	7	Canton	81

Table C.3: Regions and subregions

Country	Region definition	Regions	Subregion definition	Subregions
Czech Republic	Oblast (NUTS-2)	8	Regions (NUTS-3)	14
Denmark	Region (NUTS-2)	5	Province (NUTS-2)	11
Djibouti	Region	6	N/A	N/A
Dominican Republic	Region	10	Province	32
Ecuador	Province	25	Canton	216
El Salvador	Department	14	Municipality	255
Estonia	Group of counties (NUTS-3)	5	County (LAU-1)	15
Fiji	Group of districts	9	N/A	N/A
France	Region (NUTS-2)	25	Department (NUTS-3)	99
Georgia	Region	12	Municipality	69
Germany	State (NUTS-1)	4	District (NUTS-3)	67
Guatemala	Region	8	Department	23
Guinea-Bissau	Region	9	Sector	42
Honduras	Department	18	Municipality	269
Hungary	Planning and statistical region (NUTS-2)	7	County (NUTS-3)	20
India	State	35	District	679
Ireland	NUTS-2 Statistical Regions	2	NUTS-3 Statistical Regions	8
Jamaica	County	3	Parish	14
Kiribati	District	4	N/A	N/A
Kyrgyzstan	Region	8	District	59
Laos	Province	18	District	145
Latvia	Statistical Regions (NUTS-3)	6	District (LAU-1)	33
Liberia	County	15	District	99
Libya	District	23	N/A	N/A
Lithuania	County (NUTS-3)	10	N/A	N/A
Madagascar	Province	6	Region	22
Malawi	Region	3	District	34
Marshall Islands	Municipality	24	N/A	N/A
Mexico	State	32	Municipality	2316
Moldova	Region	5	District	35
Mongolia	Province	22	N/A	N/A
Mozambique	Province	11	District	160
Netherlands	Province (NUTS-2)	12	N/A	N/A
New Zealand	Region	16	District	86
Norway	Region	7	County (NUTS-3)	19
Pakistan	Province	3	District	91
Palau	N/A	N/A	N/A	N/A
Papua New Guinea	Province	20	District	88
Paraguay	Department	18	District	254
Peru	Region	25	Province	196
Philippines	Region	17	Province	83
Poland	Voivodeship (NUTS-2)	16	Subregions (NUTS-3)	72
Puerto Rico	Municipality	77	N/A	N/A
Saint Kitts and Nevis	Island	2	N/A	N/A
Saint Lucia	District	8	N/A	N/A
Saint Vincent and the Grenadines	District	11	N/A	N/A
Samoa	District	9	N/A	N/A
Seychelles	N/A	N/A	N/A	N/A
South Africa	Province	9	District	52
South Sudan	State	10	District	38
Sudan	State	1	District	6
Suriname	District	10	N/A	N/A
Swaziland	N/A	N/A	N/A	N/A
Sweden	National area (NUTS-2)	8	County (NUTS-3)	21
Tanzania	Region	25	District	180
Uganda	Region	4	District	118
UK	NUTS-2 Statistical Regions	39	NUTS-3 Statistical Regions	168
Ukraine	Oblast	27	Raion	626
Uruguay	Department	19	N/A	N/A
US	State	51	County	1849
Zambia	Province	10	District	103

This table shows the definition of a region used throughout this paper for every country and the number of these regions contained in the data. The table does not contain countries for which survey data was collected as sample sizes in those countries are two small for a meaningful breakdown across sub-national units. Countries for which only data from a subset of regions was collected are marked with an asterisk. For Swaziland, information on school location was not available. For Palau and Seychelles, information on the administrative divisions in which schools are located was not gathered as the total number of schools in these countries is very small.

Table C.4: GPS coordinates data sources

Country	State/Province	Data type	Collection method	Data source	Date obtained	Completeness	Shared coordinates
Antigua and Barbuda	-	Address	Download	Educational Statistical Digest 2012; Ministry of Education, Sports, Youth and Gender Affairs; Antigua and Barbuda; Retrieved from: http://www.education.gov.ag/	07.08.2017	100%	0%
Australia	-	Coordinates	Data request	Australian Curriculum, Assessment and Reporting Authority	25.04.2017	100%	$< 1\%$
Austria	Burgenland	Address	Data request	Landesschulrat fuer Burgenland	18.04.2017	100%	6%
Austria	Niederoesterreich	Address	Data request	Landesschulrat fuer Niederoesterreich	25.04.2017	100%	2%
Austria	Oberoesterreich	Address	Data request	Landesschulrat fuer Oberoesterreich	31.03.2017	100%	2%
Austria	Steiermark	Address	Download	Schulendatei Online; Bundesministerium fuer Bildung, Wissenschaft und Forschung; Retrieved from: https://www.schulen-online.at/	25.08.2017	100%	0%
Belgium	Flanders	Coordinates	Data request	Education and Training; Flemish Community of Belgium	13.12.2017	100%	$< 1\%$
Cambodia	-	Coordinates	Download	Ministry of Education, Youth and Sports; Cambodia; Retrieved from: https://opendevelopmentcambodia.net/dataset/?id=school-of-cambodia-2012	02.10.2017	100%	1%
Canada	New Brunswick	Address	Data request	Department of Education and Early Childhood Development; New Brunswick	15.06.2017	99%	0%
Canada	Ontario	Address	Data request	Ministry of Education; Ontario	20.06.2017	100%	$< 1\%$
Chile	-	Coordinates	Download	Centro de Estudios; Ministerio de Educacion; Gobierno de Chile; Retrieved from: http://centroestudios.mineduc.cl/tp_modulos/tpm_seccion/contVentana.php?cc=2179	07.11.2016	100%	2%
Czech Republic	-	Coordinates	Data request	Ministry of Education, Youth and Sports of the Czech Republic	07.04.2017	100%	1%
Denmark	-	Address	Data request	Undervisningsministeriet; Styrelsen for It og LÃring; Center for Data og Analyse; Denmark	18.09.2017	100%	$< 1\%$
Dominican Republic	-	Coordinates	Data request	Instituto Dominicano de Evaluacion e Investigacion de la Calidad Educativa	13.04.2018	99%	$< 1\%$
Ecuador	-	Coordinates	Data request	Ministerio de Educacion; Ecuador	03.03.2017	96%	2%
El Salvador	-	Coordinates	Download	Ministerio de Educacion; Republica de El Salvador; Retrieved from: https://www.mined.gob.sv/index.php/estadisticas-educativas/item/6116-bases-de-centros	11.12.2016	99%	2%
Estonia	-	Coordinates	Data request	Analysis Department; Estonian Ministry of Education and Research	24.04.2017	100%	0%
Fiji	-	Coordinates	Download	Ministry of Education; Fiji; Retrieved from: http://www.education.gov.fj/	06.03.2018	98%	$< 1\%$
France	-	Coordinates	Download	Ministere de l'Education Nationale, de l'Enseignement Superieur et de la Recherche; France; Retrieved from: https://data.education.gouv.fr/	23.01.2017	100%	1%
Germany	Hamburg	Address	Data request	Behoerde fuer Schule und Berufsbildung; Freie und Hansestadt Hamburg	15.05.2017	100%	1%

Table C.4: GPS coordinates data sources

Country	State/Province	Data type	Collection method	Data source	Date obtained	Completeness	Shared coordinates
Germany	Hessen	Address	Data request	Hessisches Kultusministerium	28.03.2017	100%	0%
Germany	Schleswig-Holstein	Address	Data request	Ministerium fuer Schule und Berufsbildung des Landes Schleswig-Holstein	03.04.2017	100%	< 1%
Germany	Thueringen	Address	Data request	Thueringer Ministerium fuer Bildung, Jugend, und Sport	20.03.2017	100%	< 1%
Guatemala	-	Coordinates	Download	Ministerio de Educacion; Guatemala; Retrieved from: http://www.mineduc.gob.gt/ie/Ministerio de Educacion; Guatemala; Retrieved from: http://www.mineduc.gob.gt/ie/	12.11.2016	47%	15%
Guinea-Bissau	-	Coordinates	Data request	Ministerio da Educacao e do Ensino Superior; Guinea-Bissau	29.09.2017	99%	< 1%
Hungary[4]	-	Address	Data request	Hungarian Central Statistical Office	06.03.2017	100%	1%
Ireland	-	Coordinates	Download	Department of Education and Skills; Ireland; Retrieved from: http://www.education.ie/en/Publications/Statistics/Data-on-Individual-Schools/Data-on-Individual-Schools.html	25.04.2018	99%	2%
Jamaica	-	Coordinates	Data request	Ministry of Education, Jamaica	24.02.2017	100%	1%
Latvia	-	Address	Data request	Ministry of Education and Science of the Republic of Latvia	21.03.2017	97%	2%
Liberia	-	Coordinates	Download	FHI360; Retrieved from: http://fhi360odk.org/kdesktoplb_2/	14.10.2017	65%	0%
Libya	-	Coordinates	Download	Libya National Schools Assessment 2012; REACH Initiative; Retrieved from: https://data.humdata.org/dataset/reach-libya-national-schools-assessment-2012	13.02.2017	100%	26%
Lithuania	-	Address	Data request	Ministry of Education and Science of the Republic of Lithuania	06.03.2017	98%	2%
Malawi	-	Coordinates	Data request	Ministry of Education; Malawi	21.03.2017	90%	0%
Marshall Islands	-	Address	Download	Education Digest 2013-2014; Ministry of Education; Republic of the Marshall Islands; Retrieved from: http://prism.spc.int/reports/education	25.11.2016	88%	37%
Mexico	-	Coordinates	Download	Sistema Nacional de Informacion de Escuelas; Secretaria de Educacion Publica; Mexico; Retrieved from: http://www.snie.sep.gob.mx/SNIESC/	02.10.2017	100%	24%
Mongolia	-	Coordinates	Data request	National Statistical Office of Mongolia	12.01.2017	80%	0%
Mozambique	-	Coordinates	Data request	Ministerio da Educacao e Desenvolvimento Humano; Republica da Mocambique	08.05.2017	95%	2%
Netherlands	-	Address	Download	Dienst Uitvoering Onderwijs; Ministerie van OCW; Retrieved from: https://www.duo.nl/open_onderwijsdata/databestanden/po/	06.02.2017	100%	< 1%
New Zealand	-	Coordinates	Download	Ministry of Education; New Zealand Government; Retrieved from: https://www.educationcounts.govt.nz/data-services/directories/list-of-nz-schools	02.10.2017	99%	0%

Table C.4: GPS coordinates data sources

Country	State/Province	Data type	Collection method	Data source	Date obtained	Completeness	Shared coordinates
Norway	-	Address	Download	Pedlex; Retrieved from: http://skoleadresser.no/	22.08.2018	96%	< 1%
Pakistan	Punjab	Coordinates	Download	School Education Department; Government of Punjab; Retrieved from: http://schoolportal.punjab.gov.pk/census/	06.10.2017	94%	5%
Pakistan	Sindh	Coordinates	Download	Education and Literacy Department; Government of Sindh; Retrieved from: http://www.rsu-sindh.gov.pk/downloads/schoolSearch.php	05.02.2018	74%	5%
Paraguay	-	Coordinates	Download	Ministerio de Educacion y Ciencias; Paraguay; Retrieved from: http://datos.mec.gov.py/data	24.01.2017	94%	< 1%
Peru	-	Coordinates	Download	Ministerio de Educacion; Peru; Retrieved from: http://sigmed.minedu.gob.pe/mapaeducativo/Ministerio de Educacion; Peru; Retrieved from: http://sigmed.minedu.gob.pe/mapaeducativo/	27.09.2017	99%	< 1%
Philippines	-	Coordinates	Download	Department of Education; Philippines; Retrieved from: https://deped.carto.com/tables/deped_school_location_with_enrolment_2014_2015/public	23.01.2018	87%	2%
Poland	-	Address	Download	Centrum Informatyczne Edukacji; Poland; Retrieved from: https://cie.men.gov.pl/sio-strona-glowna/podstawowe-informacje-dotyczce-wykazu-szko-i-placowek-owiatowych/wykaz-wg-typow/	29.01.2018	96%	5%
Puerto Rico	-	Coordinates	Download	U.S. Department of Education; National Center for Education Statistics; Common Core of Data (CCD); Public Elementary/Secondary School Universe Survey CCD School Data 2014-15; Retrieved from: http://nces.ed.gov/ccd/	10.07.2017	100%	10%
Saint Kitts and Nevis	-	Address	Download	Statistical Digest 2013-2014; St. Kitts and Nevis Ministry of Education; Retrieved from: http://www.moeskn.org/	08.12.2016	100%	0%
Saint Lucia	-	Address	Download	Education Statistical Digest 2015; Ministry of Education, Human Resource Development and Labour; Government of St. Lucia; Retrieved from: http://education.govt.lc/publications/	08.12.2016	100%	19%
Saint Vincent and the Grenadines	-	Address	Download	Education Statistical Digest of St. Vincent & the Grenadines 2014-2015; Retrieved from: http://www.education.gov.vc/education/	08.12.2016	96%	3%
Samoa	-	Address	Download	Education Statistical Digest 2015; Ministry of Education, Sports, and Culture; Samoa; Retrieved from: http://prism.spc.int/reports/education	25.11.2016	83%	13%
South Africa	-	Coordinates	Download	Education Management Information System; National Department of Basic Education; South Africa; Retrieved from: http://www.education.gov.za/Programmes/EMIS.aspx	02.11.2016	99%	1%

Table C.4: GPS coordinates data sources

Country	State/Province	Data type	Collection method	Data source	Date obtained	Completeness	Shared coordinates
South Sudan	-	Coordinates	Download	Education Management Information System; South Sudan; Retrieved from: http://www.southsudanemis.org/data	04.11.2016	81%	41%
Tanzania	-	Coordinates	Download	President's Office; Regional Administration and Local Government; The United Republic of Tanzania; Retrieved from: http://opendata.go.tz/dataset/	02.02.2018	72%	22%
Uganda	-	Coordinates	Download	Schooling Uganda; Retrieved from: https://schooling.ug/	20.10.2017	74%	0%
UK	England	Address	Download	Department for Education; England; Retrieved from: https://data.gov.uk/dataset/	09.07.2017	100%	1%
UK	Northern Ireland	Address	Download	Department for Education; Northern Ireland; Retrieved from: http://apps.education-ni.gov.uk/appinstitutes/default.aspx	07.02.2017	100%	0%
UK	Scotland	Address	Download	Scottish Government; Retrieved from: http://www.gov.scot/Topics/Statistics/Browse/School-Education/	07.02.2017	100%	2%
UK	Wales	Address	Download	Statistics Wales; Welsh Government; Retrieved from: http://gov.wales/statistics-and-research/address-list-of-schools/?lang=en	07.02.2017	100%	3%
Uruguay	-	Coordinates	Download	Administracion Nacional de Educacion Publica; Uruguay; Retrieved from: http://sig.anep.edu.uy/siganep	02.04.2018	100%	8%
US	-	Coordinates	Download	U.S. Department of Education; National Center for Education Statistics; Common Core of Data (CCD); Public Elementary/Secondary School Universe Survey CCD School Data 2014-15; Retrieved from: http://nces.ed.gov/ccd/	10.07.2017	100%	1%
Zambia	-	Coordinates	Data request	Ministry of Education; Zambia	14.06.2016	77%	3%

This table lists the data sources for the GPS coordinates of public primary schools for all countries where such data could be obtained. The last two columns indicate the share of schools for which coordinates were obtained and the share of schools with coordinates that have identical coordinates as one or more other schools.

[4] Disclaimer: Results for Hungary have been created with the use of WTorsten_cimlista_altisk_tan_ped_2015-16.xlsx Datafile prepared upon individual request by the Hungarian Central Statistical Office (www.ksh.hu). The calculations and the conclusion are the sole intellectual products of the author Torsten Figueiredo Walter.

Table C.5: PTR variation and primary schooling outcomes across countries - full census only

	Primary school completion			Primary school survival		
	(1)	(2)	(3)	(4)	(5)	(6)
National PTR	-0.00497**	-0.000815	-0.00103	-0.00545***	-0.00155	-0.00168
	(0.00200)	(0.00272)	(0.00278)	(0.00151)	(0.00210)	(0.00207)
PTR SD		-0.0116**			-0.0114**	
		(0.00535)			(0.00451)	
(PTR SD)x(GDP pc high)			-0.00833			-0.00391
			(0.00905)			(0.00662)
(PTR SD)x(GDP pc low)			-0.0114**			-0.0117**
			(0.00542)			(0.00445)
log GDP pc	0.0502**	0.0298	0.0241	0.0786***	0.0595***	0.0436*
	(0.0247)	(0.0257)	(0.0288)	(0.0202)	(0.0206)	(0.0228)
log Population	0.00227	0.00750	0.00696	-0.00879	-0.00228	-0.00344
	(0.00823)	(0.00833)	(0.00847)	(0.00695)	(0.00707)	(0.00701)
R2	0.424	0.468	0.470	0.728	0.761	0.772
N	62	62	62	52	52	52
Mean Dep. Var.	0.884	0.884	0.884	0.807	0.807	0.807
PTR SD IQR	6.061	6.061	6.061	6.059	6.059	6.059

The primary school survival rate is defined as the survival rate until the last grade of primary education. The latest available data as of 06/05/2017 is used for each country. "Log GDP pc" is the logarithm of the GDP per capita in PPP terms. "GDP pc low (high)" is a dummy variable that takes value one when the logarithm of the GDP per capita in PPP terms is below (above) the median. Outcome variables and income and population data (both 2015) are from the UNESCO Institute for Statistics and World Bank International Comparison Program Database. The national PTR is defined as the ratio of the total number of public primary school pupils over the total number of public primary school teachers in a country. The PTR SD is defined as the PTR standard deviation across all public primary schools in a country. PTR SD IQR indicates the interquartile range in the PTR SD across countries. The sample of countries is restricted to those where school census data for the entire country is available. Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C.6: PTR variation and primary schooling outcomes across countries - weighted

	Primary school completion			Primary school survival		
	(1)	(2)	(3)	(4)	(5)	(6)
National PTR	-0.00476**	-0.00127	-0.00109	-0.00556***	-0.00334*	-0.00355*
	(0.00199)	(0.00233)	(0.00237)	(0.00146)	(0.00179)	(0.00177)
PTR SD		-0.0111**			-0.00761**	
		(0.00429)			(0.00373)	
(PTR SD)X(GDP pc high)			-0.0160*			0.00110
			(0.00922)			(0.00688)
(PTR SD)X(GDP pc low)			-0.0113**			-0.00761**
			(0.00433)			(0.00368)
log GDP pc	0.0493**	0.0241	0.0304	0.0723***	0.0548***	0.0417*
	(0.0239)	(0.0249)	(0.0271)	(0.0187)	(0.0201)	(0.0217)
log Population	-0.00117	0.00557	0.00614	-0.00989	-0.00417	-0.00474
	(0.00792)	(0.00803)	(0.00813)	(0.00643)	(0.00685)	(0.00678)
R2	0.418	0.475	0.478	0.730	0.750	0.761
N	67	67	67	57	57	57
Mean Dep. Var.	0.913	0.913	0.913	0.837	0.837	0.837
PTR SD IQR	6.759	6.759	6.759	6.689	6.689	6.689

The primary school survival rate is defined as the survival rate until the last grade of primary education. The latest available data as of 06/05/2017 is used for each country. "Log GDP pc" is the logarithm of the GDP per capita in PPP terms. "GDP pc low (high)" is a dummy variable that takes value one when the logarithm of the GDP per capita in PPP terms is below (above) the median. Outcome variables and income and population data (both 2015) are from the UNESCO Institute for Statistics and World Bank International Comparison Program Database. The national PTR is defined as the ratio of the total number of public primary school pupils over the total number of public primary school teachers in a country. The PTR SD is defined as the PTR standard deviation across all public primary schools in a country. Schools are weighted by their enrollment. PTR SD IQR indicates the interquartile range in the PTR SD across countries. Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

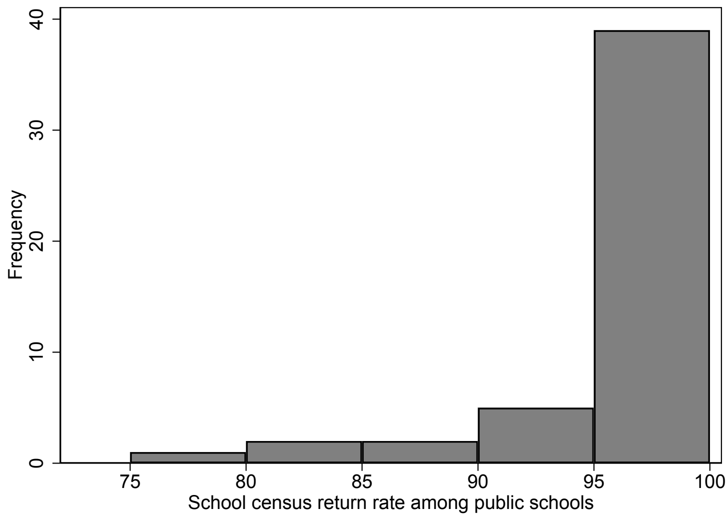


Figure C.1: School census return rates among public schools across African countries
Source: UNESCO Institute for Statistics and World Bank International Comparison Program Database. Latest available data for each country as of 13/07/2017. Sample size: 49 countries. The mean return rate across countries is 97.3%.

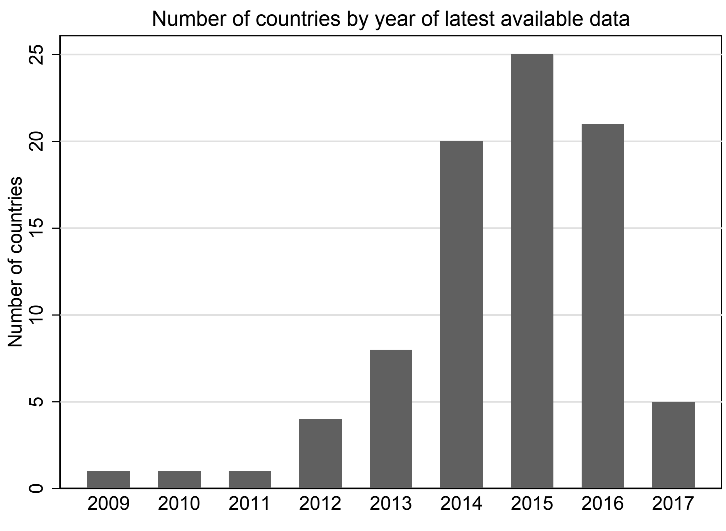


Figure C.2: Number of countries by year of data
For countries where state-level school censuses from several states were obtained the figure only contains the least recent year.

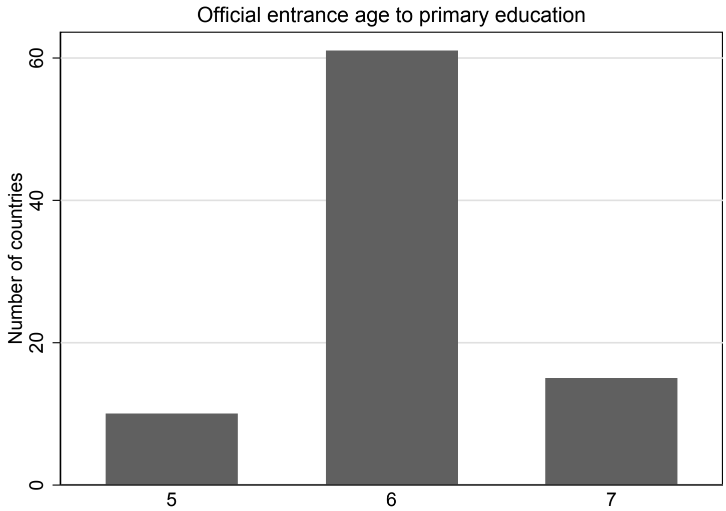


Figure C.3: Distribution of primary education entrance age across sample countries
Source: UNESCO Institute for Statistics and World Bank International Comparison Program Database. Data from 2015.

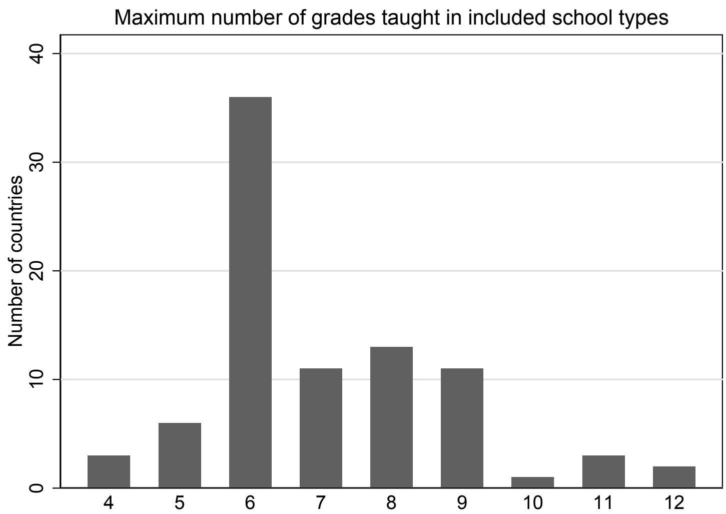


Figure C.4: Distribution of maximum number of grades taught in included school types

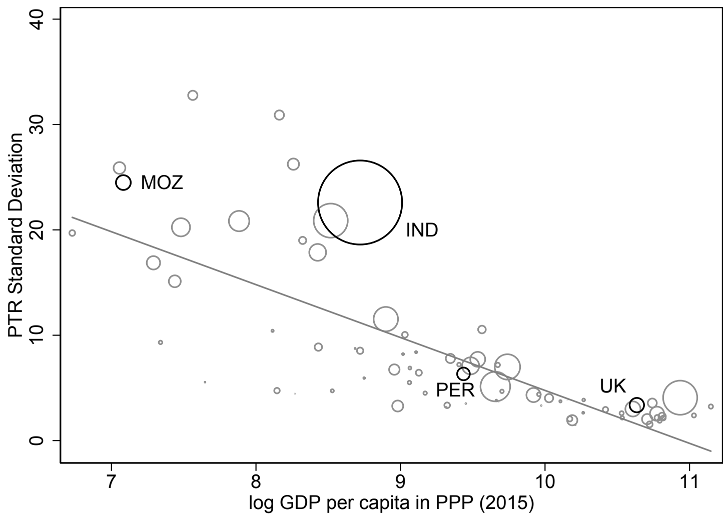


Figure C.5: PTR variation in public primary education and income across countries - weighted

The PTR standard deviation is defined as the standard deviation in PTRs across all public primary schools within a country. Schools are weighted by their enrollment. The grey line is a linear regression line. Marker size indicates the size of the primary school-aged population (ages 5 to 14) in a country. GDP per capita and population data are from the World Bank International Comparison Program database.

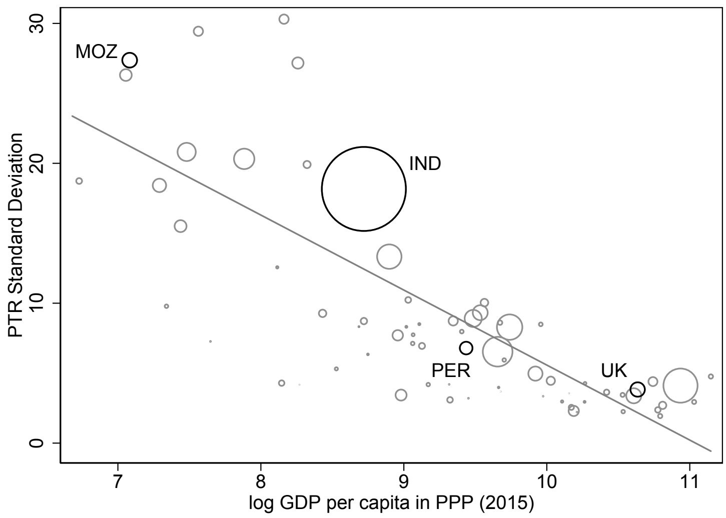


Figure C.6: PTR variation in public primary education and income across countries - full census only

The PTR standard deviation is defined as the standard deviation in PTRs across all public primary schools within a country. The grey line is a linear regression line. Marker size indicates the size of the primary school-aged population (ages 5 to 14) in a country. The sample of countries is restricted to those where school census data for the entire country is available. GDP per capita and population data are from the World Bank International Comparison Program database.

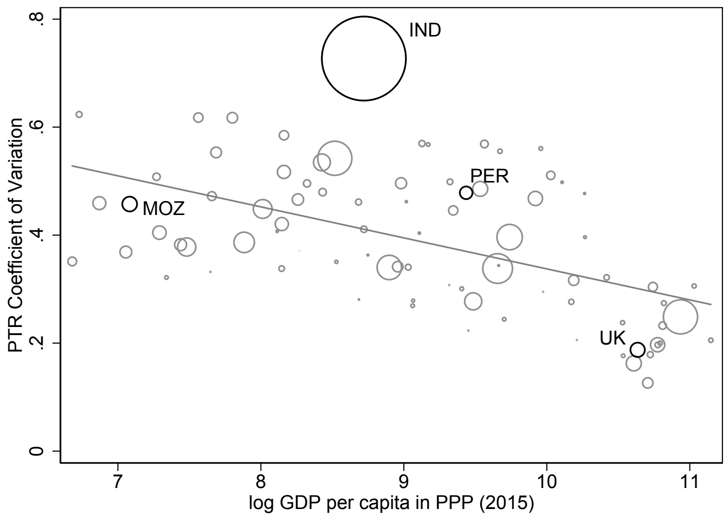


Figure C.7: PTR coefficient of variation in public primary education and income across countries
The PTR coefficient of variation is defined as the ratio of the standard deviation in PTRs across all public primary schools within a country and the mean PTR across all public primary schools within a country. Marker size indicates the size of the primary school-aged population (ages 5 to 14) in a country. GDP per capita and population data are from the World Bank International Comparison Program database.

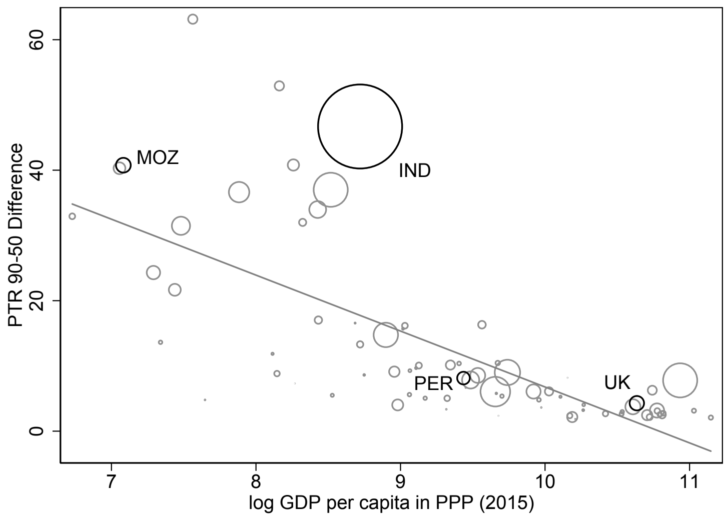


Figure C.8: Right tail of the PTR distribution and income across countries - weighted

The PTR 90-50 difference is defined as the difference between the 90th percentile and the median of the PTR distribution across all public primary schools within a country. Schools are weighted by their enrollment. Marker size indicates the size of the primary school-aged population (ages 5 to 14) in a country. GDP per capita and population data are from the World Bank International Comparison Program database.

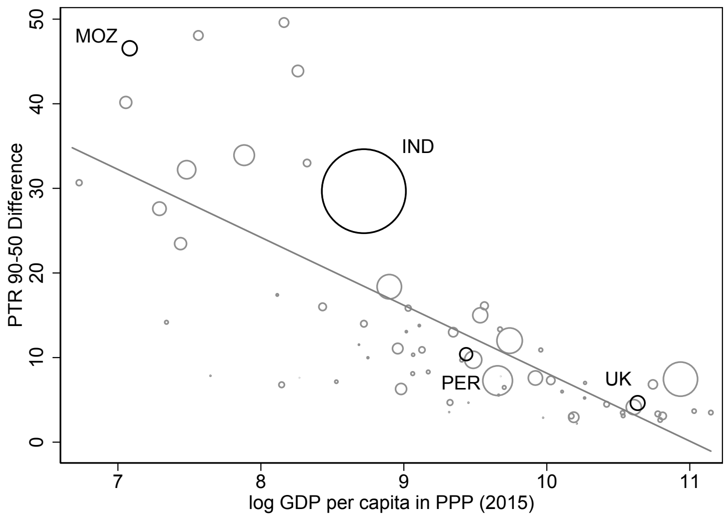


Figure C.9: Right tail of the PTR distribution and income across countries - full census only

The PTR 90-50 difference is defined as the difference between the 90th percentile and the median of the PTR distribution across all public primary schools within a country. Marker size indicates the size of the primary school-aged population (ages 5 to 14) in a country. The sample of countries is restricted to those where school census data for the entire country is available. GDP per capita and population data are from the World Bank International Comparison Program database.

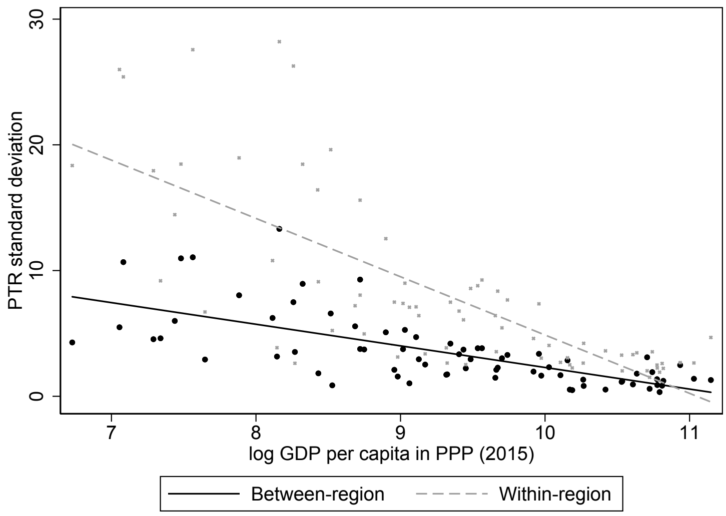


Figure C.10: Cross-school PTR variation within and between regions of a country

Regions are defined as detailed in table C.3. The sample is comprised of 72 countries. The PTR standard deviation is defined as the standard deviation in PTRs across all public primary schools within a country. Lines are linear regression lines. GDP per capita data is from the World Bank International Comparison Program database.

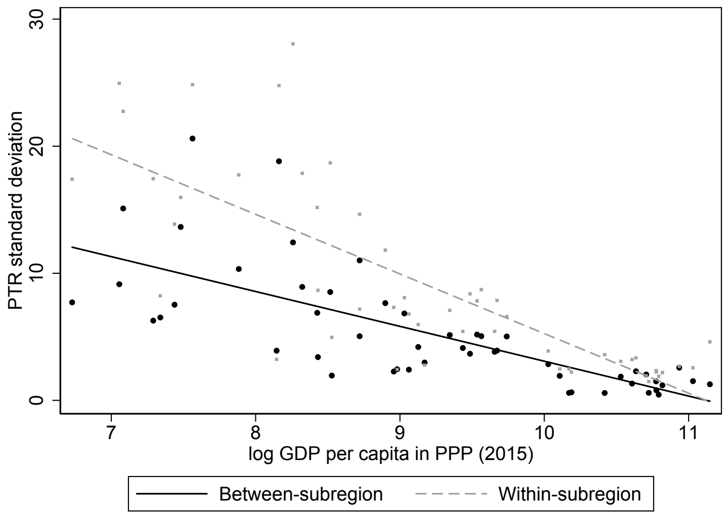


Figure C.11: Cross-school PTR variation within and between subregions of a country

Subregions are defined as detailed in table C.3. The sample is comprised of 51 countries. The PTR standard deviation is defined as the standard deviation in PTRs across all public primary schools within a country. Lines are linear regression lines. GDP per capita data is from the World Bank International Comparison Program database.

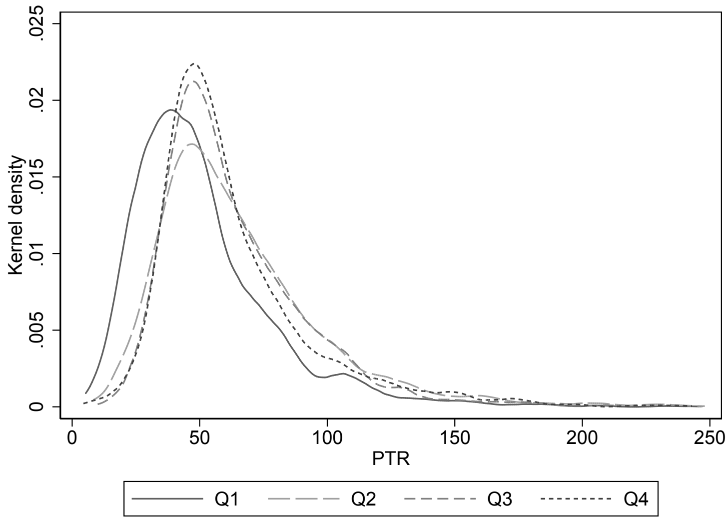


Figure C.12: PTR distribution by quartile of population density (GPW) in Mozambique
Population density at each public primary school is measured as the density within a circle of 3km radius around the school according to data from the Gridded Population of the World (v4).

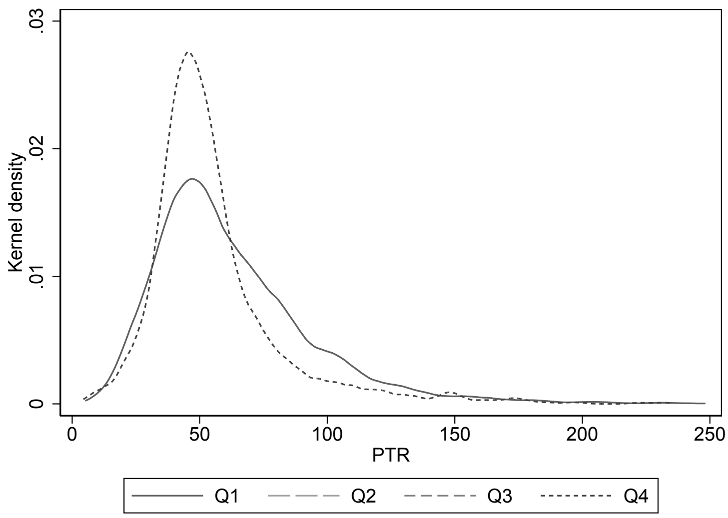


Figure C.13: PTR distribution by quartile of nighttime luminosity in Mozambique
Nighttime luminosity data is from the Earth Observation Group, NOAA National Geophysical Data Center (VIIRS 2015).

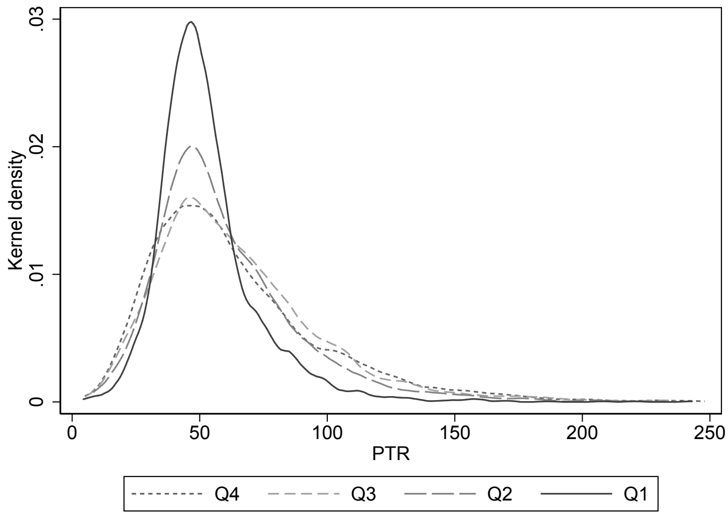
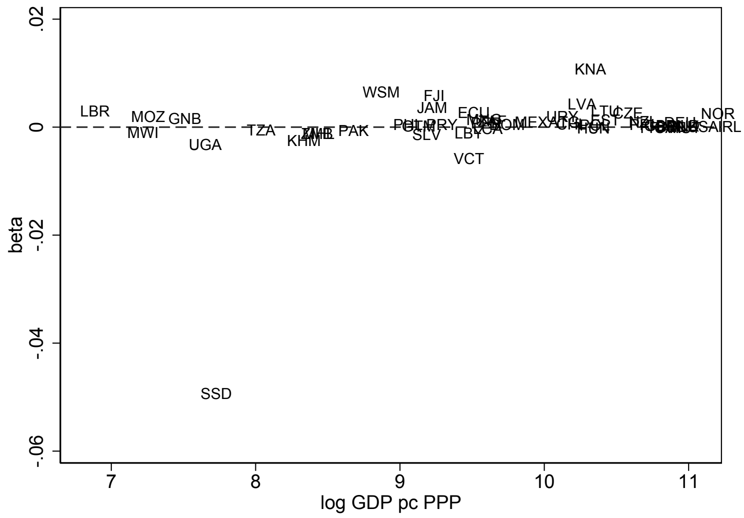
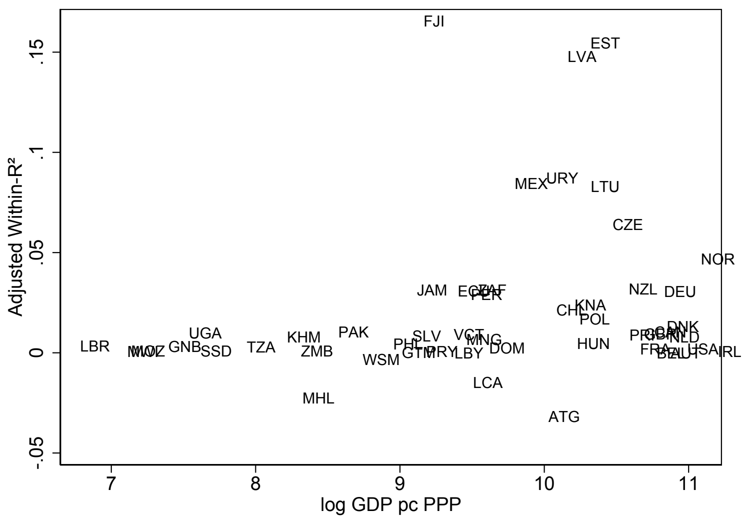


Figure C.14: PTR distribution by quartile of travel time to the closest city in Mozambique

Travel time to closest city is taken from the accessibility to cities data set from the Malaria Atlas Project at Oxford University (Weiss et al. 2018).



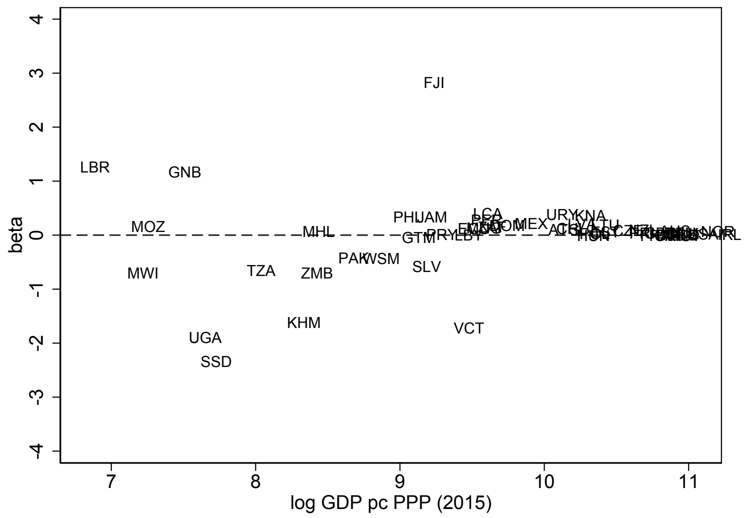
(a) Regression coefficient



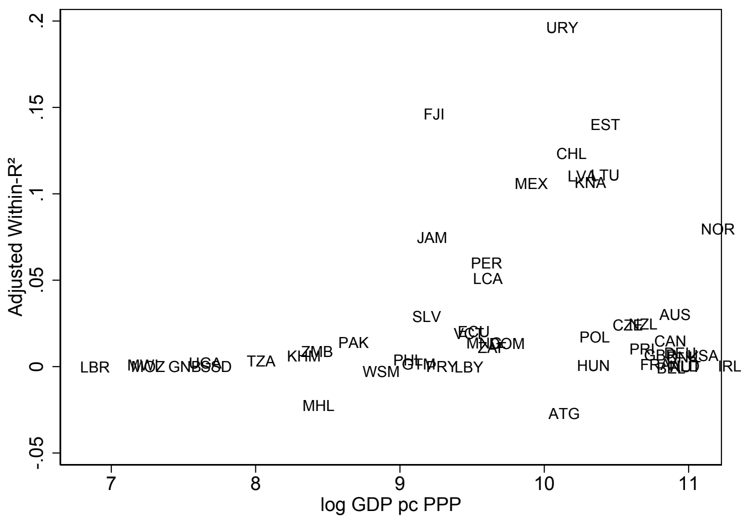
(b) Within-region R^2

Figure C.15: PTRs and population density (GPW) across countries by income

Beta is the regression coefficient from a country-specific school-level regression of PTR on population density within a circle of 3km around the school as given by the Gridded Population of the World (v4) data, controlling for region fixed effects. The adjusted within-region R^2 is from the same regression. Regions are defined as detailed in table C.3. The sample is restricted to 51 countries for which school coordinates were obtained. GDP per capita data is from the World Bank International Comparison Program database.



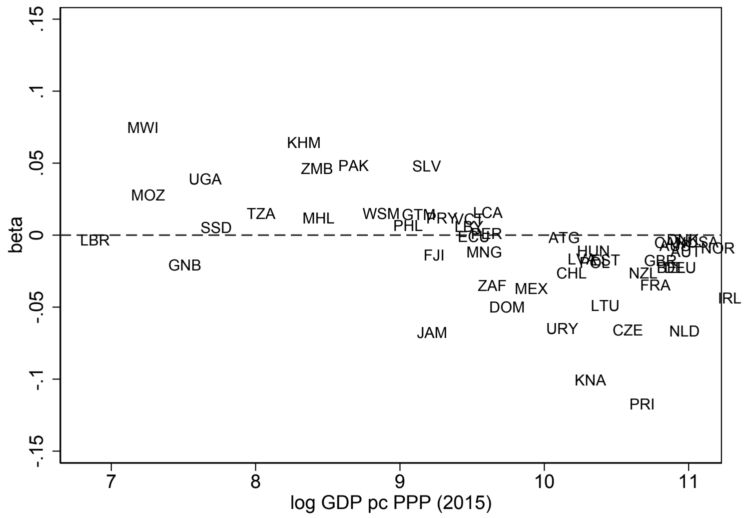
(a) Regression coefficient



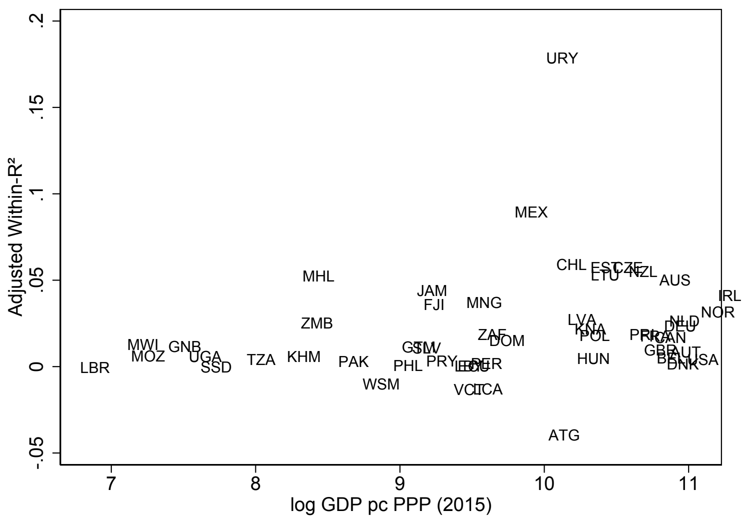
(b) Within-region R^2

Figure C.16: PTRs and nighttime luminosity across countries by income

Beta is the regression coefficient from a country-specific school-level regression of PTR on average nighttime luminosity within a circle of 3km radius around the school in 2015, controlling for region fixed effects. The adjusted within-region R^2 *is from the same regression. Nighttime luminosity was obtained from VIIRS nighttime lights data from the Earth Observation Group, NOAA National Geophysical Data Center. Regions are defined as detailed in table C.3. The sample is restricted to 50 countries for which school coordinates were obtained. GDP per capita data is from the World Bank International Comparison Program database.*



(a) Regression coefficient



(b) Within-region R^2

Figure C.17: PTRs and travel time to closest city across countries by income

Beta is the regression coefficient from a country-specific school-level regression of PTR on travel time to closest city as given by Weiss et al. (2018), controlling for region fixed effects. The adjusted within-region R^2 is from the same regression. Regions are defined as detailed in table C.3. The sample is restricted to 50 countries for which school coordinates were obtained. GDP per capita data is from the World Bank International Comparison Program database (2015).

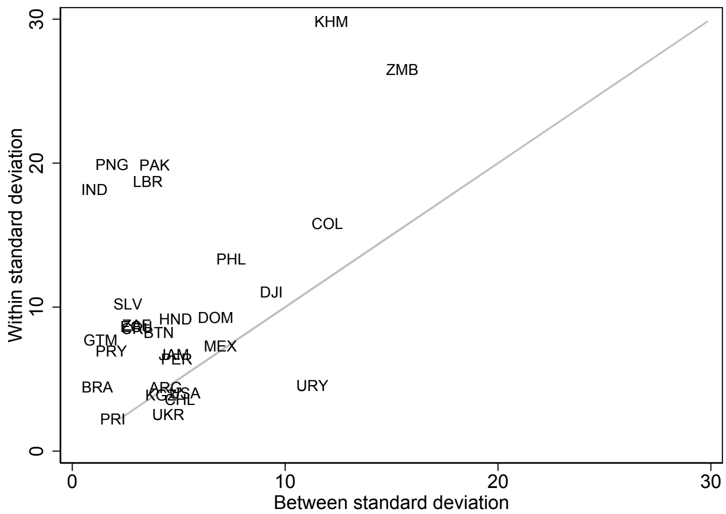


Figure C.18: PTR variation within and between rural/urban classification
Rural/urban indicator is country-specific, as provided by the school census data.

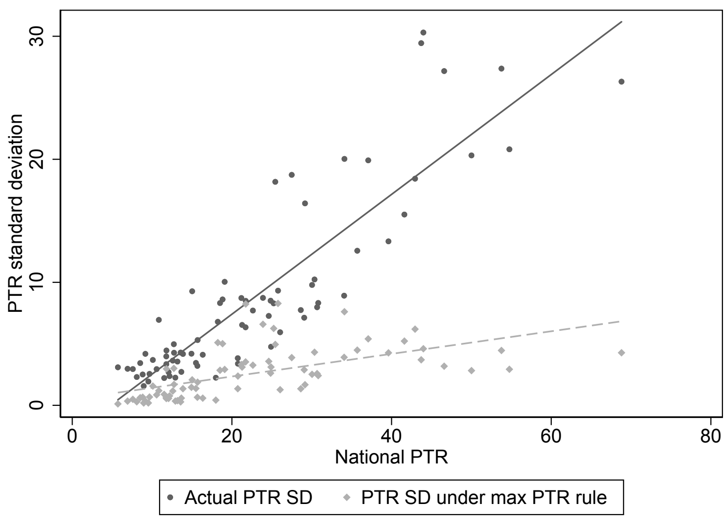


Figure C.19: Actual and counterfactual PTR standard deviation and national PTR across countries

Appendix D

Appendix to Chapter 4: Gains from Teacher Reallocation?

Table D.1: Data sources: primary school teacher salaries

Country	Salary (GDP pc)	Year	Source
Cambodia	2.15	2017	Sokhean, B.; Sineat, Y.; Amaro, Y. Teachers' wage rally blocked. In Phnom Penh Post (06/10/2017). Retrieved from: *https://www.phnompenhpost.com/national/teachers-wage-rally-blocked* on 23/08/2018.
Guinea-Bissau	4.4	2006	World Bank (2015). Education Public Expenditure Review in Zambia.
India	3.0	2004/05	Dreze, Jean; Sen, Amartya. 2013. An uncertain glory: India and its contradictions. Princeton, NJ: Princeton University Press.
Malawi	6.3	2008	World Bank (2015). Education Public Expenditure Review in Zambia.
Mozambique	4.0	2003	World Bank (2015). Education Public Expenditure Review in Zambia.
Tanzania	3.8	2004	UNESCO Pole de Dakar (2009). The teacher challenge - Universal primary education in Africa. Dakar: UNESCO-BREDA.
Zambia	6.7	2014	World Bank (2015). Education Public Expenditure Review in Zambia.

This table indicates the teacher salaries used for costing the results of the counterfactual simulations in chapter 4 and their sources.

Table D.2: Subnational units for counterfactual simulation

Country	Subnational unit
Argentina	Province
Cambodia	Province
Cape Verde	Country
Chile	Province

Table D.2: Subnational units for counterfactual simulation

Country	**Subnational unit**
Colombia	Department
Djibouti	Region
El Salvador	Department
Guinea-Bissau	Region
Honduras	Department
India	District
Laos	Province
Malawi	District
Mozambique	District
Peru	Province
Saint Lucia	Country
Saint Vincent	Country
Sweden	NUTS-2
Tanzania	District
UK (England)	NUTS-1
Zambia	District

This table indicates the subnational unit used for the simulation of the third counterfactual, optimal allocation of teachers within subnational units, for each country.

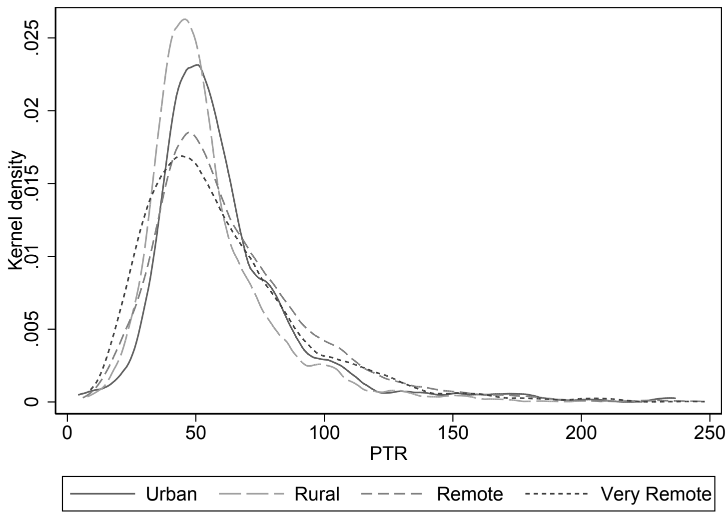


Figure D.1: Distribution of PTRs by hardship allowance category in Mozambique
Data sources: Education Management Information System, Ministerio de Educacao e Desenvolvimento Humano, Mozambque (2016) and Conselho de Ministros, Decreto n° 91/2009.

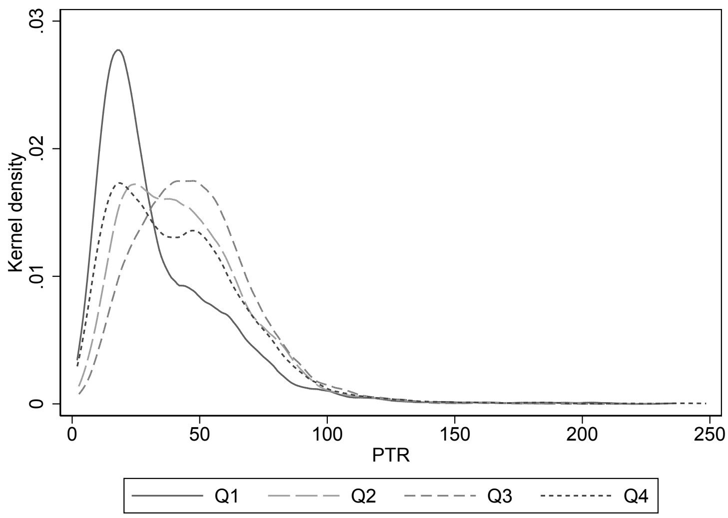


Figure D.2: Distribution of PTRs by quartile of teacher retention rates in Uganda
Data source: Education Management Information System, Ministry of Education, Uganda (2016).

Appendix E

Appendix to Chapter 6: Human resource misallocation in other public sectors? Evidence from the staffing of Zambian and English primary care facilities

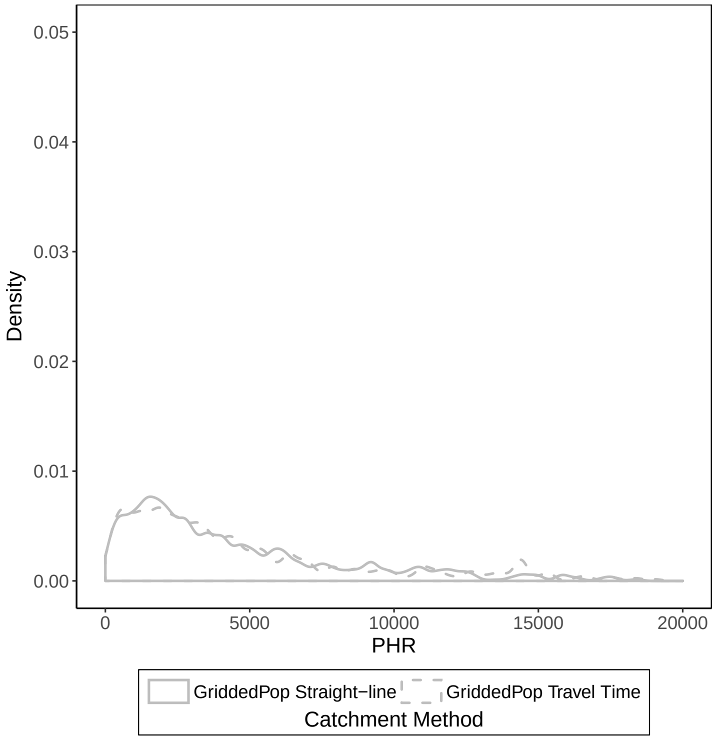


Figure E.1: Distribution of PHRs across primary care facilities in Zambia - including facilities without official catchment population counts
Sample comprised of all primary care facilities with GPS coordinates and staffing data. Health workers include all medical staff. Facilities are weighted by catchment population.

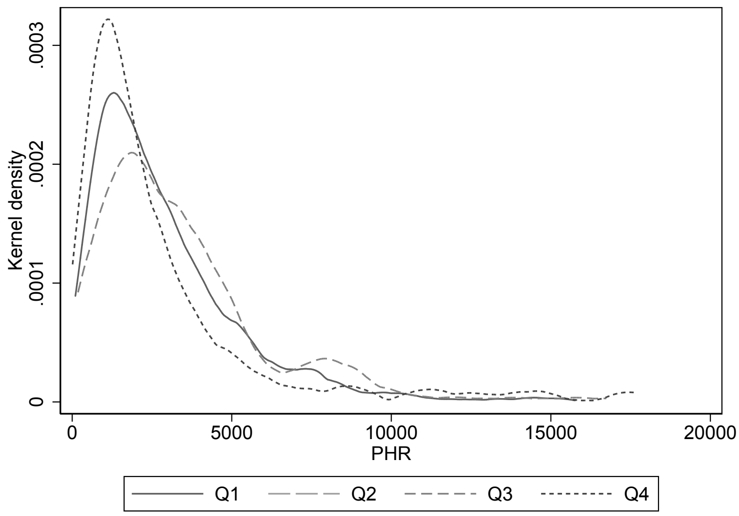


Figure E.2: PHR distribution by quartile of nighttime luminosity in Zambia

Nighttime luminosity at a health facility is defined as mean luminosity within a circle of 3km radius around the facility. Data is from the Earth Observation Group, NOAA National Geophysical Data Center (VIIRS 2015).

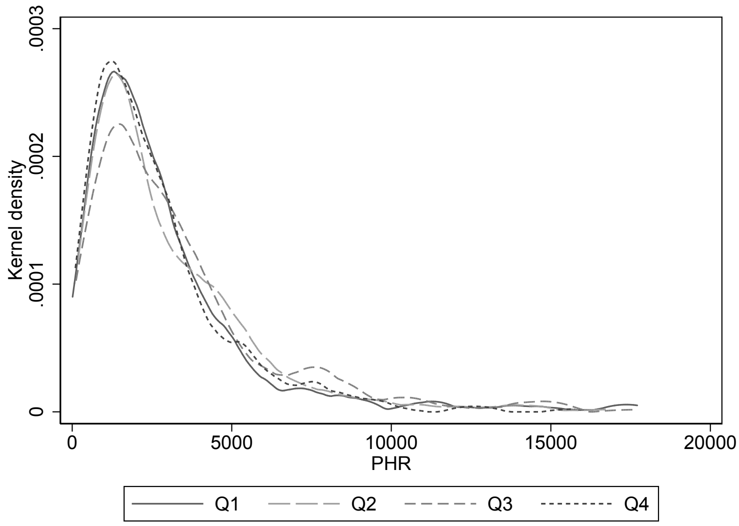


Figure E.3: PHR distribution by quartile of travel time to closest city in Zambia

Travel time to closest city is taken from the accessibility to cities data set from the Malaria Atlas Project at Oxford University (Weiss et al. 2018).